

Optimizing vendor selection in a two-stage outsourcing process

Qing Cao^{a,*}, Qiong Wang^b

^a*Bloch School of Business and Public Administration, University of Missouri at Kansas City, Kansas City, Missouri 64110, USA*

^b*Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974, USA*

Available online 3 March 2006

Abstract

The decision processes surrounding outsourcing are complicated by the very nature of uncertainty involved in the outsourcing process and by poor vendor management. In this study, we focus on vendor selection, one of the two basic issues of vendor management in outsourcing. Due to the limitation of the classic one-stage vendor selection model, we propose a two-stage vendor selection research framework in outsourcing. The first stage is a trial phase that helps the client to find the best match between the vendor and the outsourced project. In the second stage, the client employs the chosen vendor for the full implementation of the project. We formulate this selection decision under the two-stage framework as a combinatorial optimization model. We analyze the complexity of the problem and develop a solution procedure to find the exact optimal solution. By applying this model to numerical case studies, we demonstrate that benefit to adopt two-stage process to the vendor depends on information improvement in the first stage and the client's ability to adapt to updated knowledge. We also argue that the selection of vendors for the first stage testing is more about creating a good vendor portfolio than simply picking the frontrunners.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Outsourcing; Vendor selection; Combinatorial Optimization; Risk management

1. Introduction

Outsourcing is concerned with the transferring of internal production functions of goods or services to an external provider [1]. The outsourcing trend continues unabated in the US despite the negative politicization of outsourcing by the media and politicians. For instance, Lacity and Willcocks [2] report that IT outsourcing contracts alone expect to reach US\$ 156 billion in 2004. It is also estimated that more than 50 percent of companies in the US will outsource their IT functions in 2006 [3].

Moreover, the practice of outsourcing is expanding in both scope and sophistication. It used to be the case that outsourcing only involves non-core activities with the purpose of helping firms to reduce costs and concentrate on their core competency. Nowadays outsourcing has become a universal phenomenon in every area of business, such as engineering, research and development (R&D), new product development, and marketing [4]. Outsourcing also plays a strategic role by helping firms to acquire new capabilities, to bring about fundamental changes to managerial strategy and organizational structure, and to facilitate the transformation of business models [5]. Many firms are looking to outsourcing for a variety of benefits, from cost savings to increased flexibility, and from improvement in service quality to better access to the state-of-the-art technology [6].

* Corresponding author. Tel.: +1 816 235 6242; fax: +1 816 235 6506.

E-mail address: caoq@umkc.edu (Q. Cao).

Nevertheless, not every outsourcing project is a success story. Reported problems include the degradation of service [7], the lack of cost reduction [8], and disagreements between clients and vendors [9]. While some failures can be explained by the complexity and uncertainty involved in the outsourcing process, others are due to the poor management of vendors.

Vendor management involves both vendor selection and client–vendor relationship control. Many researchers have addressed the latter issue by the employment of transaction cost theory [7,10] or the analysis of incomplete contracts [11]. The objective is for the client to induce the optimal performance from the vendors. It is equally important to select the right vendor for the task. As noted by Power et al. [12], one pitfall for outsourcing failure is the minimal knowledge of outsourcing methodologies, especially in the vendor selection process. This issue has been addressed by many, but the majority of the work focuses on the *single-stage* process; that is, to build mathematical models to help the client to choose one vendor from many *solely* based on the past performance data. Such approaches imply that vendors' past performance guarantees future results, an assumption does not necessarily apply because of ever changing technology, a high degree of heterogeneity of outsourced projects, intrinsic variation, and low predictability in vendor performance [13].

Recognizing the limitation of the single-stage approach, several researchers have discussed a two-stage process [13,14]. The first stage is a trial phase that helps the client to find the best match between the vendor and the outsourced project. In the second stage, the client employs the chosen vendor for the full implementation of the project. The dominant theme of the existing work is on the dynamics of client–vendor interaction in the first stage, analyzed under the game theory framework.

In this paper, we subscribe to the notion of the two-stage process, but try to answer a different question: under resource constraints, how should the client determine the scope of the first-stage testing and decide which vendors should participate based on their prior performance information? The fundamental tradeoff of this problem is embedded in the allocation of a fixed budget between the two stages. To improve its chance of getting a high-performing vendor for the task, the client would like to invite as more candidates to participate in the first-stage testing as possible. However, involving too many vendors depletes the client's budget and consequently, limits its ability to carry out the project to a full extent in the second stage. We argue that the first-stage testing is a process of gathering vendor information, so the investment in the process should depend on the value of the information to be gathered. These values depend on both the quality of information itself (by how much the client can expect to improve its knowledge about the vendors by the first stage testing), and the client's ability to act on that information. We also argue that choosing participating vendors in the first-stage testing resembles portfolio management, i.e., vendors should be selected based on not only their own virtue, but also the complement of their performance profiles.

We formulate the vendor selection problem as a combinatorial optimization model and prove its NP-hardness by showing the model can be reduced to a standard knapsack problem. Therefore, no efficient solution procedure is known, nor are we attempting to find one in this paper. We develop a search procedure that guarantees an exact optimal solution, augmented by some test conditions that eliminates as many enumerations as possible. Even though in the worst case, the algorithm cannot avoid a complete enumeration, we bank on the fact that in normal situations, the number of potential vendors for a client is quite limited in which case the algorithm should be appropriate.

The rest of this paper is organized as follows. We review related work in Section 2, formulate the combinatorial optimization model for vendor selection in Section 3, and discuss model properties and solution procedure in Section 4. In Section 5, we present a numerical example that illustrates many interesting insights. Section 6 concludes the paper.

2. Literature review

Vendor selection is an important aspect of many outsourcing projects. Back in the late 1960s, Wind and Robinson [15] developed the commonly-used linear weighting method that ranks and selects vendors based on their performance on multiple measurements. Later researches expanded this classical approach along several dimensions. For instance, Gregory [16] introduced matrix representation of data into this scheme and rated different vendors for their quota allocations. Monozka and Trecha [17] proposed multiple criteria vendor service factor ratings and an overall supplier performance index.

In comparison with the linear weighting methods, mathematical programming offers a more effective tool to handle the vendor selection problem because of the latter's ability to explicitly optimize some stated objective [18]. Commonly

used models include linear programming (LP), mixed integer programming (MLP), analytical hierarchical process (AHP), data envelopment analysis (DEA), and goal programming (GP) [19–23]. Methodologies vary, but so does the problem setting. For example, Murthy et al. [19] developed an LP model for make-to-order items in a situation that involves fixed cost, shared capacity constraints, and volume-based discounts. The objective is to minimize total sourcing and purchasing costs. Dahel [20] studied a multi-vendor, multi-product, and competitive sourcing environment. They developed a mixed integer programming model to determine both the number of vendors and the allocation of total orders to each of them. The AHP method is used in Handfield et al. [21] to generate weights for vendor selection problem. Ghodspour and O'Brien [22] offered an integrated decision support system that combines the AHP method with LP.

There are many other studies that focus on the multi-attribute aspect of the vendor selection problem (price, quality, flexibility, and delivery time); for instance, the negotiation model in Zhu [23], DEA models by Liu et al. [24] and Weber et al. [25], GP models in Sharma et al. [26], and fuzzy mixed integer goal programming model in Kumar et al. [27]. Nevertheless, these models are commonly formulated as an one-stage decision problem, and suffer from the assumption that the client has perfect information about the vendors, a luxury it does not have in the real-world.

To address the uncertainty in the vendor selection problem, Snir and Hitt [13] presented a two-stage contracting framework and analyzed it from the game theory perspective. They argued that the model creates many benefits for the client, especially when they have no prior information about vendors (so all vendors are equally qualified) and vendors have to invest in the first stage (so there is little cost to the client, and vendors have a financial stake in the success of the projects).

In this paper, we study the two-stage process for a different situation. We assume that the client has prior information about the vendors. The client wants to improve that knowledge by engaging multiple vendors in a pilot study in the first stage before making a full commitment to one of them in the second stage. We also consider the client has a fixed budget and it has to share a substantial amount of the first-stage cost. This gives rise to a different problem that we intend to address.

3. Vendor selection as a combinatorial optimization problem

Consider a client who contemplates a two-stage approach for an outsourcing project: a pilot study in the first stage to select a vendor for the task, and the full implementation of the project in the second stage by the chosen vendor. The client has a fixed budget C for the entire project and has to divide the fund into c_1 and $c_2 \equiv C - c_1$ for the respective use in the first and second stages.

The benefit of the project is derived from the second-stage implementation, which depends on both the implementation budget $C - c_1$ and the performance of the selected vendor. Assume that vendor performance can be rated by a single measurement with L levels, indexed by $l = 1, 2, \dots, L$.

Let $b_l(c_2)$ be the benefit of the project that can be realized if the vendor performance is at level l and the implementation budget is at c_2 . Let the performance index be sorted in an ascending order so that

$$b_l(c_2) \geq b_{l'}(c_2) \quad \text{if } l > l'.$$

Also assume $\partial b_l(c_2)/\partial c_2 \geq 0$, i.e., the client's benefit from the project increases with the budget size.

Let there be n vendors, indexed by $i = 1, 2, \dots, n$. Before the first-stage testing, the client may evaluate a vendor based on its past performance. The result of that evaluation can be characterized as a prior distribution over the performance level, i.e. q_{il} ($i = 1, 2, \dots, n$ and $l = 1, 2, \dots, L$), interpreted as the client's subjective probability that vendor i will give a l level performance:

$$\sum_{l=1}^L q_{il} = 1, \quad i = 1, 2, \dots, n.$$

The client gets to know how well a vendor will perform on the given project only after subjecting the latter to the first-stage testing.

Let s_i denote the amount of money that the client needs to spend on vendor i for the first-stage testing, and let x_i ($i = 1, 2, \dots, n$) be a set of binary variables, where $x_i = 1$ if vendor i is selected for the first-stage testing and $x_i = 0$ otherwise. Then the client’s first stage expenditure is

$$c_1 = \sum_{i=1}^n s_i x_i. \tag{1}$$

In selecting these vendors, the client faces a tradeoff. To enhance its chance to recruit a top-performing vendor for the second-stage implementation, the client should involve as many vendors in the testing as possible, which means more expenditure in the first-stage. However, to derive the maximum benefit from implementing the project, more money should be left for the second stage implementation.

To quantify this tradeoff, let $Q_{il} = q_{i1} + q_{i2} + \dots + q_{il}$ ($i = 1, 2, \dots$) be the probability that vendor i ’s performance is below level l . It follows that $\prod_{i=1}^n Q_{il}^{x_i}$ is the probability that the best performance of all tested vendors does not exceed level l and

$$G_l(x_1, x_2, \dots, x_n) = \prod_{i=1}^n Q_{il}^{x_i} - \prod_{i=1}^n Q_{i,l-1}^{x_i} \tag{2}$$

is probability that level l is the best performance of tested vendors, and thus is the performance that the client will receive in the second stage. Given these probabilities, the client’s expected benefit is

$$\sum_{l=1}^L G_l(x_1, x_2, \dots, x_n) b_L (C - c_1).$$

Apply Eqs. (1) and (2) and observe that $Q_{iL} = 1$ for all $i = 1, 2, \dots, n$ (i.e., no vendor can perform better than the best level L), the client’s decision of selecting vendors for testing can be formulated as the following combinatorial optimization problem:

$$\begin{aligned} B &= \max_{x_1, x_2, \dots, x_n} \left\{ \prod_{i=1}^n Q_{i1}^{x_i} b_1 (C - c_1) + \sum_{l=2}^L \left(\prod_{i=1}^n Q_{il}^{x_i} - \prod_{i=1}^n Q_{i,l-1}^{x_i} \right) b_l (C - c_1) \right\} \\ &\quad \left. \sum_{i=1}^n s_i x_i = c_1, \sum_{i=1}^n x_i \geq 2, x_i = 0, 1 \right\} \\ &= \max_{x_1, x_2, \dots, x_n} \left\{ b_1 (C - c_1) + \sum_{l=2}^L \Delta b_l (C - c_1) \prod_{i=1}^n Q_{il}^{x_i} \mid \sum_{i=1}^n s_i x_i = c_1, \sum_{i=1}^n x_i \geq 2, x_i = 0, 1 \right\}, \tag{3} \end{aligned}$$

where $\Delta b_l (C - c_1) = b_l (C - c_1) - b_{l-1} (C - c_1)$. The constraint $\sum_{i=1}^n x_i \geq 2$ is imposed because there is no reason to conduct a first-stage testing with only one vendor. By doing so, the client would have effectively committed this only vendor for the project since whatever result turned out by the testing, the client would not have a second choice. In this case, it is better to skip the first stage all together and save the money for the second stage implementation.

4. Model property and solution procedure

As we will show in this section, the combinatorial problem formulated in (3) can be reduced to a standard knapsack model, and thus is NP-complete, which is known to be hard to solve [28]. In cases like this, one has to choose between the efficiency of the algorithm and the accuracy of the solution, and we will select the latter. We develop a search procedure that always finds the exact optimal solution but in the worst case, requires evaluation of every potential solution. We argue that for our application, the client usually has a limited number of vendors to choose from. Consequently, even the complete enumeration of all solutions is within easy reach of available computing power. We also try to improve the efficiency of the algorithm by imposing test conditions that can eliminate unnecessary enumerations.

We start from the following proposition on the hardness of the model.

Proposition 1. *There exist instances of model (3) that is equivalent to a knapsack problem, so that optimization problem in (3) is NP-complete.*

Proof. Consider a special case in which $L = 2$ and $b_l(C - c_1)$ ($l = 1, 2$) take the forms of the following step function:

$$b_l(C - c_1) = \begin{cases} \bar{b}_l & \text{if } C - c_1 \geq C_0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $0 < C_0 < C$. To get a positive benefit in this case, the choice of x_i ($i = 1, 2, \dots, n$) is subject to

$$\sum_{i=1}^n s_i x_i \leq C - C_0.$$

Consequently the objective function in (3) becomes

$$\bar{b}_2 - (\bar{b}_2 - \bar{b}_1) \prod_{i=1}^n Q_{i1}^{x_i}$$

and model (3) becomes equivalent to

$$\min_{x_1, x_2, \dots, x_n} \left\{ \prod_{i=1}^n Q_{i1}^{x_i} \mid \sum_{i=1}^n s_i x_i \leq C - C_0, \sum_{i=1}^n x_i \geq 2, x_i = 0, 1 \right\}. \tag{5}$$

For any non-trivial case where $0 < Q_{i1} \leq 1^1$ ($i = 1, 2, \dots, n$), let $a_i = -\log(Q_{i1}) \geq 0$. Consequently, (5) is the same as

$$\max_{x_1, x_2, \dots, x_n} \left\{ \sum_{i=1}^n a_i x_i \mid \sum_{i=1}^n s_i x_i \leq C - C_0, \sum_{i=1}^n x_i \geq 2, x_i = 0, 1 \right\}. \tag{6}$$

If (6) is not NP-complete, then neither would be the problem

$$\max_{x_1, x_2, \dots, x_n} \left\{ \sum_{i=1}^n a_i x_i \mid \sum_{i=1}^n s_i x_i \leq C - C_0, x_i = 0, 1 \right\}, \tag{7}$$

which is clearly not true because (7) is the standard formulation of the knapsack problem [28]. \square

Given the NP-completeness of our problem, developing an efficient procedure to find the exact optimal solution is impossible (at least for now). Therefore, in this paper, we develop a depth-first search procedure that guarantees the exact solution at the expense of computation time. To improve computational efficiency, we introduce some test conditions that enable us to reduce the number of potential solutions to be searched. These conditions are based on the following two observations.

First, if vendor i dominates vendor j in both performance and testing cost, i.e.,

$$Q_{il} \leq Q_{jl} \quad \text{for all } l \text{ and } s_i \leq s_j,$$

then it is never optimal to select vendor j but not i for the first-stage testing.

Second, suppose we know there is a feasible solution that gives B_0 as the value of the objective function. Suppose that $\{x_{i_1}, x_{i_2}, \dots, x_{i_n}\}$ is a set of selected vendors. Let

$$c_1 = s_{i_1} x_{i_1} + s_{i_2} x_{i_2} + \dots + s_{i_n} x_{i_n} \quad \text{and} \quad \underline{s} = \min\{s_i : i \neq i_1, i_2, \dots, i_n\}.$$

If $B_0 \geq f_L(C - c_1 - \underline{s})$, then even testing an additional vendor guarantees the best performance for the client, the improvement in performance is out-weighted by the impact of additional funding reduction for the second stage. Therefore, we can safely exclude the solution that includes $\{x_{i_1}, x_{i_2}, \dots, x_{i_n}\}$ and any other vendor.

¹ The case where there exists i' that $Q_{i'1} = 0$ (so that $q_{i'2} = 1$) is trivial because it means the best performance is guaranteed by vendor i' , in which scenario there is no need for the first stage testing.

To give a step-by-step description of our search algorithm, define T as a first-in-last-out (FILO) stack that keeps the indexes of vendors that are currently chosen. For given T , let $B(T)$ be the corresponding value of the objective function. During the search process, vendors are dynamically added to or removed from T , which leads to changes of $B(T)$. Define T^* as the best selection that has been found so far in the process, and D as the set of vendors that should not be excluded from the selection based on previous searching result.

The process includes the following steps:

1. Initialize T , T^* , and D to $\{\}$ and let $i_0 = 1$, where i_0 is the starting point of the search and will be updated in the process.
2. If $i_0 = n$, we have reached the end of our algorithm. T^* is the optimal selection, and $B(T^*)$ is the value of the optimal solution.
3. Let $i_1 = i_0$.
4. Add i_1 to T and evaluate the current objective function:

$$B(T) = b_1(C - c_1) + \sum_{l=2}^L \Delta b_l(C - c_1) \prod_{i \in T} Q_{il}, \quad \text{where } c_1 = \sum_{i \in T} s_i.$$

Also let $\underline{s} = \min_{i \notin T} s_i$.

5. Compare the current solution with the best solution found so far

if $B(T) > B(T^*)$, then let $T^* = T$.

6. Consider changing T
 - (a) If $(i_1 < n)$ AND vendor $(i_1 + 1 \notin D)$ AND $(b_L(C - \sum_{i \in T} s_i - \underline{s}) < B(T^*))$, then let $i_1 = i_1 + 1$ and go to step 4.
 - (b) Else if $i_1 = n$, let $i_1 = i'$, where i' is the index of the penultimate element in T (recall T is a FILO stack). Remove the last two vendors from T and proceed to step 7.
 - (c) Else if vendor $i_1 + 1 \notin D$, then let $i_1 = i_1 + 1$ and go back to the beginning of step 6.
 - (d) Else if $b_L(C - \sum_{i \in T} s_i - \underline{s}) > B(T^*)$, then remove vendor i_1 from T , let $i_1 = i_1 + 1$, and go to step 4.
7. If $i_1 > i_0$, then go to step 4. Otherwise, add vendor i_0 to D .

Add any vendor $i > i_0$ to D if $Q_{i0l} \leq Q_{il}$ for all l and $s_i s_{i0} \leq s_i$.

Let $i_0 = i_0 + 1$ and proceed to step 2.

5. Numerical examples

5.1. Base case

In the base case, we consider 10 vendors (indexed from 1 to 10) and 4 scenarios. Each scenario is attached with a score: 1 (“poor”), 2 (“fair”), 3 (“good”), and 4 (“excellent”). Table 1 gives prior distribution of vendors’ performance over these scenarios that are known to the client before the first-stage testing. Also included is the client’s cost of going through the first stage with vendor i ($i = 1, 2, \dots, 10$).

Vendors 1–3 are high-performance firms who command a premium in the testing cost. Vendors 9 and 10 reside on the other side of the spectrum. The rest fall into the middle ranges in both performance and cost, which can be further characterized as stable performers (vendors 4–6) and volatile ones (vendors 7 and 8).

The client’s benefit from the project is formulated as

$$b_l = A_l(C - c_1)^\varepsilon, \tag{8}$$

where A_l ($l = 1, 2, 3, 4$) and ε are inputs. Recall C is the client’s total budget and $C - c_1$ is the amount of budget left for implementation after the first-stage testing. The value of ε is in the range of $(0, 1)$, reflecting a positive contribution

Table 1
Performance profile and testing cost

| Vender | Performance probability | | | | Cost |
|--------|-------------------------|------|------|-----------|------|
| | Poor | Fair | Good | Excellent | |
| 1 | 0.1 | 0.0 | 0.3 | 0.6 | 25 |
| 2 | 0.2 | 0.1 | 0.0 | 0.7 | 25 |
| 3 | 0.0 | 0.0 | 0.5 | 0.5 | 25 |
| 4 | 0.0 | 0.1 | 0.9 | 0.0 | 15 |
| 5 | 0.0 | 0.2 | 0.7 | 0.1 | 15 |
| 6 | 0.0 | 0.1 | 0.8 | 0.1 | 15 |
| 7 | 0.4 | 0.0 | 0.0 | 0.6 | 15 |
| 8 | 0.3 | 0.2 | 0.0 | 0.5 | 15 |
| 9 | 0.1 | 0.5 | 0.4 | 0.0 | 5 |
| 10 | 0.1 | 0.6 | 0.3 | 0.0 | 5 |

Table 2
Result: distribution of payoff over different scenarios

| Scenario <i>l</i> | Benefit <i>b_l</i> | Probability <i>q_{il}</i> (%) | Expected value <i>b_lq_{il}</i> |
|----------------------|---------------------------------|--|--|
| 1 (poor) | 0.0 | 0.0 | 0.0 |
| 2 (fair) | 312.8 | 0.0 | 0.0 |
| 3 (good) | 521.4 | 6 | 31.3 |
| 4 (excellent) | 730.0 | 94.0 | 686.2 |
| Total | | 100 | 717.5 |

of more implementation spending to the benefit of project, as well as the diminishing return of that contribution. In this example, we let $\varepsilon = 0.5$. The value of A_l increases with l , i.e., the benefit of the project is larger with a better performing vendor. In the example, we let $A_1 = 0$, $A_2 = 15$, $A_3 = 25$, and $A_4 = 35$.

Suppose the client has a total budget of \$500.00 million. Given the above inputs, the model produces the following outcome:

1. The client should spend 65 (more than 10% of total budget) on the first stage to screen vendors. The vendors selected are 2, 3, and 7, i.e., two high-cost, high-performance firms, and one firm with medium cost and volatile performance.
2. The client’s expected payoff is \$717.50 million, the break down of which is given in Table 2.

5.2. Discussions and managerial implications

5.2.1. When the screening process is helpful

The first issue we address with this study is when the two-stage process is a better approach than the first-stage alternative. In the latter case, the client picks the vendor based on prior information on performance distribution (Table 1) and starts directly with the implementation of the project. By doing so, the client avoids the testing cost and puts the money into the direct generation of the benefits.

The value of the first-stage testing comes from the improvement of the client’s understanding of the vendors and the wiser selection of a vendor for project implementation. Correspondingly, whether this process is worth paying for depends on two factors. First, to what extent the client can improve its knowledge about the vendors by testing and second, how much flexibility does the client have to adapt its selection decision based on updated information from the testing. From this perspective, one would naturally expect the following results.

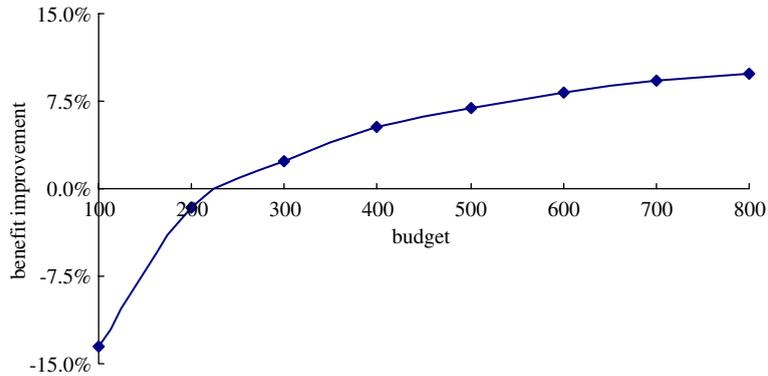


Fig. 1. Change of relative benefit improvement with budget.

Consider the base case under the one-stage process. From Table 1, the average performance of a vendor can be ranked by

$$\bar{A}_i = \sum_{l=1}^4 A_l q_{il}, \quad i = 1, 2, \dots, 10, \tag{9}$$

where the values of A_l and q_{il} are given in Section 4.1 given above. Since

$$\bar{A}_3 = \max\{\bar{A}_i : i = 1, 2, \dots, 10\},$$

one would naturally choose vendor 3. The expected benefit is then

$$B' = \bar{A}_3 C^e = 30\sqrt{500} = 670.5. \tag{10}$$

Recall from Table 2, that for the same case, the expected benefit of the two-stage process is 717.5, 7% improvement in benefit. Nevertheless, this advantage may not necessarily persist in all situations.

We conduct a few numerical experiments with the client’s budget. We vary the budget size from \$100 to \$800 million in increments of \$100 million. The relative improvement of the second stage process is defined by

$$\frac{B - B'}{B'}, \tag{11}$$

where B is the expected benefit of the two-stage process given by and B' is that from the one-stage process, calculated similarly as in Eq. (10) (with different values for budget C). Fig. 1 shows reduction (even negative) of the improvement by the two-stage process as the client’s budget shrinks. The reason for this reduction of the advantage is that with a smaller budget, the client can only recruit a smaller number of vendors for the first-stage testing. Therefore, what it can do after getting better information about vendors is more limited. For instance, when the budget gets as small as \$100 million, the client’s benefit is maximized if it recruits only one vendor, vendor 3, for the first-stage testing (if we drop the requirement that $\sum_{i=1}^n x_i \geq 2$). In this case, the vendor selection decision is not contingent on the outcome of the first-stage, and the vendor is better off with the one-stage approach.

Another important factor for accessing the value of the two-stage process is the improvement of vendor information. Suppose we keep the \$500 million budget constant and vary the difference between scenarios 3 and 4 by letting

$$A_3 = A_0 - \delta, \quad A_4 = A_0 + \delta.$$

In the base case, $A_3 = 25$ and $A_4 = 35$ so $A_0 = 30$ and $\delta = 5$. Let A_0 be a constant and δ fluctuates. Therefore, the average performance of our preferred vendor, vendor 3, does not change, but the uncertainty about actual performance goes up and down with the value of δ as δ goes up and down. Fig. 2 shows the change of relative improvement (see Eq. (11)) in the expected benefits achieved by the two-stage process with the changing value of uncertainty parameter, δ .

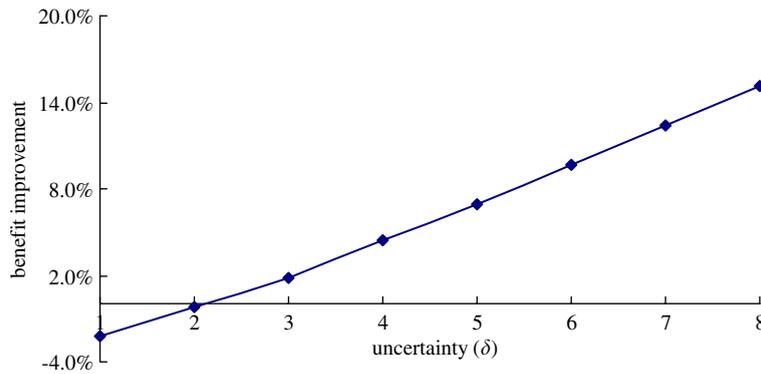


Fig. 2. Change of relative benefit improvement with uncertainty.

Table 3
Ranking of the vendors

| Vendor | Ave. perf. (\bar{A}_i) | Screening cost |
|--------|----------------------------|----------------|
| 1 | 28.5 | 25 |
| 2 | 26 | 25 |
| 3 | 30 | 25 |
| 4 | 24 | 15 |
| 5 | 24 | 15 |
| 6 | 25 | 15 |
| 7 | 21 | 15 |
| 8 | 20.5 | 15 |
| 9 | 17.5 | 5 |
| 10 | 16.4 | 5 |

For a small value of δ (low uncertainty), the little information improvements the client gains from the first-stage testing does not justify the cost of going through the process, which is an effect exemplified by the negative value of the relative benefit improvement. The gain becomes substantial when δ becomes large.

The above analysis has the following managerial implications. In deciding whether or not to adopt a two-stage process, a manager has to ask herself two questions:

1. Is there enough uncertainty about the vendors that can be resolved by first-stage testing?
2. Do I have enough budgetary flexibility to make the best vendor selection based on the outcome of the first-stage testing?

Only when the answers to both questions are positive should the two-stage process take place.

5.2.2. Which vendors to choose

Suppose it makes economic sense to adopt a two-stage process. Then the next question is: which vendors should the client choose to be included in the first-stage testing? We argue that this initial selection decision is not a simple process of picking frontrunners, but rather a process of creating a well-balanced candidate portfolio. The following numerical examples support this argument.

Let us first determine each vendor’s average performance, \bar{A}_i (as in Eq. (9)) and give these numbers in Table 3. The table shows that vendors 4–6 score better than vendor 7 in the average performance while the screen cost are the same for the four vendors. Despite these inputs, in the optimal solution, the chosen vendor should be 7 (together with vendors 2 and 3) instead of other vendors 4–6.

This seeming contradiction is actually a reflection of the basic principle of managing a portfolio, i.e., a candidate should be chosen based on not only its individual merit, but also how well it complement other candidates’ performance.

Table 4
Optimal choice of vendors for different payoffs

| Vendor selected | Case 1 ($A_4 = 35$) | | | Case 2 ($A_4 = 30$) | | |
|-----------------|-----------------------|-------------|----------------|-----------------------|-------------|----------------|
| | Probability (%) | Performance | Expected value | Probability (%) | Performance | Expected value |
| Poor | 0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Fair | 0.4 | 313 | 1.3 | 1.2 | 316 | 3.9 |
| Good | 3.6 | 521 | 18.8 | 9.6 | 527 | 50.6 |
| Excel. | 95.2 | 730 | 694.9 | 89.4 | 632 | 564.5 |
| Total | 100.0 | | 715 | 100 | | 619 |

In this case, the existence of vendor 3 as a candidate guarantees that the client will not end up with the first (poor) and second (fair) scenarios.² Therefore, the only remaining concern is whether the last scenario (excellent performance by at least one candidate vendor) will emerge from the screening process. On this front, selecting vendor 7 contributes more to maximize the chance of that happening than choosing any of other three vendors.

The tradeoff between reward and risk is another defining feature of portfolio management. A decision-maker is willing to take more risk of failure if doing so would result in a higher reward in success. This tendency is readily observable in our problem in the observable choice of vendors.

Consider the base case in which the optimal choices of vendors are 2, 3, and 7. Here, vendor 3 is an attractive candidate from both risk and reward perspective. Its performance is no worse than “good”, eliminating the client’s risk of ending up with a poor or fair performing vendor. It also has a good chance of being an “excellent” vendor, giving the client a good reward.

Now suppose we eliminate vendor 3 as a possible candidate by attaching a very high testing cost to it, which vendor should be chosen as its replacement? Table 3 shows an interesting phenomenon. In case 1, we keep everything else unchanged from the base case except making vendor 3 un-selectable, then another high-cost, high-performance vendor, vendor 1, is chosen in vendor 3’s place. Given there is a small probability for vendor 1 to perform poorly, the client will be taking a slight risk of failure rate of 0.4%.

Case 2 is the same as Case 1 except we reduce the value A_4 from 35 to 30, i.e., in the best scenario, the benefit of the project is

$$30(C - c_1)^{0.5} \text{ instead of } 35(C - c_1)^{0.5}.$$

This change effectively reduces the client’s reward from having an “excellent” vendor. As a consequence, the client is no longer willing to take the risk. As shown in the table, instead of vendor 1, the client selects vendor 6, a decision made to eliminate the chance of the worst case, but one that also has a smaller probability to achieve the best case.

We put these two cases side by side in Table 4. Summarizing from the data, we conclude that in selecting vendors for the first-stage testing, the right combination is more important than picking the best performer. The balance between risk and reward is of vital importance.

6. Conclusions and future research

The focus of our research is to help clients to render vendor selection decisions in the outsourcing process. The proposed two-stage vendor selection model is more adequate than the conventional single-stage model, the limitations of which were previously discussed. Our two-stage vendor selection model assumes that the client has prior information about the vendors and that the client wants to improve that knowledge by engaging multiple vendors in a pilot study in the first stage before making a full commitment to one of them in the second stage. It also considers that the client has a fixed budget and it has to share a substantial amount of the first-stage cost.

² The worst the client can do is to pick vendor 3 after testing, whose performance will be either good (scenario 3) or excellent (scenario 4), but cannot be bad, fair, or poor.

6.1. Modeling contributions

Our main contribution in this area is to formalize the vendor selection decision as a combinatorial optimization problem. We proved that the problem is NP-complete by showing it can be reduced to a canonical Knapsack problem. We provide a calculation procedure that delivers the exact optimal solution and includes features that improve computational efficiency.

6.2. Limitations

Our conclusions may be affected by certain limitations in this study, such as the use of the simulated data for case studies. We also assumed that vendors' performances are measured by a single criterion. In practice, multi-criteria measurements are often used in the vendor selection processes.

6.3. Managerial implications

The first managerial implication of the proposed two-stage vendor selection model is that it can be used as part of an easy-to-use decision support system, which is in great need for outsourcing practices [12]. Second, the existing single-stage vendor selection approach is more restrictive in application value thanks to its assumptions, which imply that vendors' past performance guarantees future results than our vendor selection model. A third managerial implication is that this research provides a generalized research framework, which can be used to analyze numerous outsourcing scenarios. Finally, the proposed vendor selection model cannot only answer the question in terms of which vendor the clients should choose but also help clients to determine whether it is worthwhile to adopt a two-stage outsourcing process.

6.4. Future research

The vendor selection problem in this paper does not incorporate the tradeoff between insourcing and outsourcing. This would render an already complex problem more computationally intractable. Nevertheless, it is an important problem for future research. Future research needs to apply real world performance probability data or at least come up with a proper performance probability distribution in the analysis. It will also be interesting to conduct cross-project comparison and test the validity of the proposed vendor selection model.

References

- [1] Domberger S, Jensen PH, Stonecash RE. Examining the magnitude and sources of cost savings associated with outsourcing. *Public Performance and Management Review* 2002;26(2):148.
- [2] Lacity MC, Willcocks LP. An empirical investigation of information technology sourcing practices: lessons from experience. *MIS Quarterly* 1998;22(3):363–408.
- [3] King WR. Outsourcing and the future of IT. *Information Systems Management* 2004;21(4):83–4.
- [4] Gottfredson M, Puryear R, Phillips S. Strategic sourcing: from periphery to the core. *Harvard Business Review* 2005;83(2):132–9.
- [5] Linder JC. Transformational outsourcing. *MIT Sloan Management Review* 2004;45(2):52–8.
- [6] McFarlan FW, Nolan RL. How to manage an IT outsourcing alliance. *Sloan Management Review* 1995;36(2):9–23.
- [7] Kern T, Willcocks L, van Heck E. The winner's curse in IT outsourcing: strategies for avoiding relational trauma. *California Management Review* 2002;44(2):47–69.
- [8] Scheier RL. Outsourcing's fine print. *Computerworld* 1996;30(3):70.
- [9] Earl MJ. The risks of outsourcing IT. *Sloan Management Review* 1996;37(3):26–32.
- [10] Willcocks L, Lacity M, Kern T. Risk mitigation in IT outsourcing strategy revisited: longitudinal case research at LISA. *Journal of Strategic Information Systems* 1999;8(3):285–314.
- [11] Nam K, Rajagopalan S, Rao HR, Chaudhury A. A two-level investigation of information systems outsourcing. *Communications of the ACM* 1996;39(7):36–44.
- [12] Power M, Bonifazi C, Desouza KC. The ten outsourcing traps to avoid. *The Journal of Business Strategy* 2004;25(2):37–42.
- [13] Snir EM, Hitt LM. Vendor screening in information technology contracting with a pilot project. *Journal of Organizational Computing and Electronic Commerce* 2004;14(1):61–88.
- [14] DiRomualdo A, Gurbaxani V. Strategic intent for IT outsourcing. *Sloan Management Review* 1998;39(4):67–80.
- [15] Wind Y, Robinson PJ. The determinants of vendor selection: the evaluation function approach. *Journal of Purchasing and Materials Management* 1968; 29–41.

- [16] Gregory RE. Source selection: a matrix approach. *Journal of Purchasing and Materials Management* 1986; 24–9.
- [17] Monozka RM, Trecha SJ. Cost-based supplier performance evaluation. *Journal of Purchasing and Materials Management* 1988; 2–7.
- [18] Degraeve Z, Roodhooft F, van Doveren B. The use of total cost of ownership for strategic procurement: a company-wide management information system. *The Journal of the Operational Research Society* 2005;56(1):51–9.
- [19] Murthy NN, Soni S, Ghosh S. A framework for facilitating sourcing and allocation decisions for make-to-order items. *Decision Sciences* 2004;35(4):609–27.
- [20] Dahel NE. Vendor selection and order quantity allocation in volume discount environments. *Supply Chain Management* 2003;8(3/4):335–42.
- [21] Handfield RB, Ragatz GL, Petersen KJ, Monczka RM. Involving suppliers in new product development. *California Management Review* 1999;42(1):59–72.
- [22] Ghodsypour SH, O'Brien C. The total cost of logistics in supplier selection, under conditions of multiple sourcing, multiple criteria and capacity constraint. *International Journal of Production Economics* 2001;73(1):15–27.
- [23] Zhu J. A buyer-seller game model for selection and negotiation of purchasing bids: extensions and new models. *European Journal of Operational Research* 2004;154(1):150–6.
- [24] Liu F, Ding FY, Lall V. Using data envelopment analysis to compare suppliers for supplier selection and performance improvement. *Supply Chain Management: An International Journal* 2000;5(3):143–50.
- [25] Weber CA, Current JR, Desai A. An optimization approach to determining the number of vendors to employ. *Supply Chain Management: An International Journal* 2000;2(5):90–8.
- [26] Sharma D, Benton WC, Srivastava R. Competitive strategy and purchasing decisions. In: *Proceedings of the 1989 annual conference of the decision sciences institute*; 1989. p. 1088–90.
- [27] Kumar M, Vrat P, Shankar R. A fuzzy goal programming approach for vendor selection problem in a supply chain. *Computers and Industrial Engineering* 2004;46(1):69–85.
- [28] Carey MR, Johnson DS. *Computer and intractability, a guide to the theory of NP completeness*. New York: W. H. Freeman; 1979.