Optimal Selection of Line Extensions: Incorporating Operational, Financial, and Marketing Constraints

Liying Mu, Milind Dawande, Srinagesh Gavirneni, and Chelliah Sriskandarajah

***Abstract*—Line extensions—variants of existing products with new appearances, functions, or forms—constitute a significant fraction of products launched each year. While line extensions typically share components with existing products, they can also cannibalize the demand of existing products. The choice of an appropriate set of line extensions is an important recurring issue for a firm and they must incorporate constraints that arise naturally from operational, financial, and marketing considerations, resulting in an analytically challenging optimization problem. We consider several variants of this problem, develop efficient heuristics that deliver near-optimal solutions, and derive a variety of interesting insights on the inherent tradeoffs in the selection of line extensions.**

***Index Terms*—Budget constraint, cardinality constraint, line extensions, polynomial-time algorithms.**

I. Introduction

FIRMS ACROSS industries face an increasing need to offer a greater level of the product variety that more closely meets the diverse needs of their customers in global markets. Over the years, a preferred route to offer the required product variety has been through line extensions: variants of existing products with new appearances, functions, or forms. Popular examples of line extensions include iPhone 4 (an extension of iPhone 3GS), Audi A8 (an extension of Audi A6), and collectors' editions of DVDs (extensions of their standard titles). For consumer goods, an overwhelming majority of new products launched each year are line extensions (see, e.g., [1] and [2]).

For a typical multicomponent product, the number of *feasible* line extensions can be high. For example, for an automobile, there could be hundreds of potential line extensions based on variations in external design, internal design, trim design, and chassis design. Other products where the possible number of line extensions can be high include televisions and computers. The decision to introduce a specific set of line extensions is jointly driven by constraints that originate from operational, financial, and marketing considerations.

Manuscript received January 20, 2013; revised April 29, 2013, November 11, 2013, April 3, 2014, and April 6, 2014; accepted May 23, 2014. Date of publication June 25, 2014; date of current version October 16, 2014. Review of this manuscript was arranged by Department Editor B. Jiang.

L. Mu and M. Dawande are with the Naveen Jindal School of Management, University of Texas at Dallas, Dallas, TX 75080 USA (e-mail: muliying@utdallas.edu; milind@utdallas.edu).

S. Gavirneni is with the Johnson School of Management, Cornell University, Ithaca, NY 14850 USA (e-mail: nagesh@cornell.edu).

C. Sriskandarajah is with Mays Business School, Texas A & M University, College Station, TX 77843 USA (e-mail: chelliah@mays.tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEM.2014.2327484

Operational Constraints: An extension typically shares several components with (one or more) existing products. Depending on the extent of this sharing and the nature of the new components (if any) required, an extension could incur a lower or a higher development cost as compared to a new product, and could enjoy economies of scale in the production or purchase of the components that it shares with existing products (see [3]–[5]). A case in point is the manufacturing of automobiles: Toyota introduced the Sienna mini-van as a line extension of its basic Camry sedan. Both Sienna and Camry are assembled on the same platform and share many underbody parts (see [6], [7]).

Financial and Extension-Overload Considerations: When selecting the set of product extensions to introduce, firms are typically restricted by budget available for investment. In some situations, however, the dominant consideration could be the harmful impact of introducing too many extensions (see, e.g., [8]–[12]). These restrictions can manifest themselves as a budget constraint on the total expense (see [13], [14]) or an upper bound on the cardinality of the set of chosen line extensions ([15], [16]).

Marketing Constraints: From a marketing perspective, it is important to consider the impact new line extensions have on the market share of existing products. This relationship, known as cannibalization, has been extensively discussed in the marketing literature (see [17]–[19]).

From the set of potential line extensions, it is desirable for a firm to choose a subset that maximizes profit subject to the operational, financial, and marketing constraints. This is a challenging combinatorial optimization problem, for which an exhaustive enumeration approach is not practical. We build on the basic model proposed by Ramdas and Sawhney [20], and analyze two important variants of the basic model: 1) a budget-constrained version; and 2) a cardinality-constrained version. We resolve the complexities of both variants and develop efficient heuristics that deliver near-optimal solutions. Using these heuristics, we derive several interesting managerial insights on the impact of problem-specific characteristics on the optimal fraction of line extensions (i.e., the ratio of the number of extensions chosen to the total number of possible extensions). For practicing managers, the recommendations from our analysis can be summarized as follows.

1) When determining a firm's new product development strategy, an important consideration is the relationship between investment needed prior to product launch and the costs that will be incurred during manufacturing and distribution. The prelaunch investment could consist of the product design, component development, supplier

identification and collaboration, and integration testing. The postlaunch costs could consist of material and labor costs, and the costs associated with logistics and distribution. Our results show that for industries in which prelaunch costs are significantly larger than postlaunch costs (e.g., computers, smart phones, and digital cameras), the number of new product introductions will be fewer. On the other hand, for industries in which the postlaunch costs constitute a dominant portion (e.g., apparel, housewares, and stationary), the number of new product introductions will be larger.

2) The degree of commonality among the components plays a key role in identifying the number of new products that a firm will introduce. When the commonality increases, then for the same set of components, the number of possible product designs is larger, which in turn implies that the fixed costs associated with new product development can be distributed over a larger volume of sales. Thus, as commonality among the components increases, the number of new product introductions increases as well.

3) In addition to the product design and component development, supply chain contracting is an important constituent of a product development strategy. A fundamental aspect of a supply chain contract is the negotiation of the volume discount strategy. Two possibilities are marginal discounts—after a threshold volume, each additional unit is discounted; and all-unit discounts—after a threshold volume, all the units are discounted but the undiscounted price is higher than that for a marginal discount policy. We recommend that when the product–component matrix is sparse, the firm should negotiate a marginal discount supply contract. As the product–component matrix becomes denser the marginal discount strategy becomes less preferable, the extent of which depends on the price difference between the two contracting schemes.

4) The capital budget made available to new product development or the limit on the number of new product introductions are strategic decisions that should be made by the firm in order to maximize its return on investment (ROI). Our results show that, as the budget or the cardinality constraint is relaxed, the ROI first increases rapidly and then levels off. There exists a budget limit or a cardinality level that is a sweet spot in this relationship (which is shaped like a knee joint) and the firm should use that level in order to maximize its ROI.

The remainder of this paper is organized as follows. Section II provides an overview of the related work in the operations and marketing literature. In Section III, we introduce the basic model of [20] and then discuss an efficiently solvable special case. We also develop an algorithm with an attractive performance bound for the general version of the model. Section IV discusses the budget-constrained variant of the basic model, while the cardinality-constrained variant is presented in Appendix B. In Section V, we examine the performance of our heuristics based on a test bed of realistic instances. Section VI summarizes useful insights on the optimal set of line extensions and concludes the paper.

II. Review of the Related Literature

Modularity and Commonality have long been used as the main strategies (see, e.g., [21], [22]) for balancing operational, financial, and marketing considerations in product-line selection. As mentioned in [23], the extensive academic debate on these strategies has, however, resulted in a proliferation of views on product modularity that have rendered its essential definition ambiguous. To address this issue, Salvador [23] develops a product modularity construct that accurately accounts for the presence of both component separability and component combinability. These two concepts have largely evolved independently; e.g., Gupta and Krishnan [24] and Sanderson and Uzumeri [25] focus on component sharing, while Meyer and Lehnerd [26] and Meyer and Utterback [27] focus on modularity. Mikkola and Gassmann [28] develop a modularization function that mathematically captures the degree of modularity in a given product architecture. Blecker and Abdelkafi [29] develop a total-commonality index that, in a step-by-step manner, captures the overall commonality of a product family. Agrawal *et al.* [30] study the optimal product-line design by simultaneously utilizing component sharing and modularity, and demonstrates that, in some settings, the strategy of using only component sharing is better than that of integrating both component sharing and modularity. Our paper builds on this observation and focuses on a system in which component sharing is the predominant strategy in the product-line design. Agrawal *et al.* [30] illustrate a solution procedure on a two-product example and recommends the extrapolation to a larger product family as an avenue for future research. Our study is motivated by this objective.

Most of the aforementioned research on component commonality in the design of line extensions has side-stepped the difficulties associated with allocating component costs (development and procurement) among all the products that use a component. Labro [31] examines the issues associated with cost accounting methods used to allocate costs in the presence of common components and concludes that there is significant disagreement on whether using common components increases or decreases costs, thus highlighting the need for further research. Jiao *et al.* [32] observe that existing cost accounting methods, when used to allocate fixed and variable costs across multiple products, may result in distorted allocations. The authors identify the need for well-defined performance measures that can enable companies to better manage their product and component portfolios. This confusion on the relationship between cost and degree of component commonality motivates the use of rigorous mathematical models and sophisticated optimization algorithms to make decisions. Ramdas and Sawhney [20] formulate an optimization problem for choosing a set of line extensions, by incorporating the revenue interactions between new products and the cost interactions due to the sharing of components. As mentioned earlier, we build on the work in [20] by analyzing budget- and cardinality-constrained versions of the model studied in that paper. Jans *et al.* [33] formulate a mixed-integer nonlinear optimization model to determine the optimal assignment of components to products, but do not offer generalizable managerial insights due to their focus on a single industrial setting.

Sanchez and Sudharshan [34] tabulate various limitations—time, cost, inaccuracies—of traditional market research methods and postulate that rapid introduction of new products along with flexible manufacturing systems can enable firms to conduct real-time market research. The authors recognize that reuse of existing components and using the same component in multiple products can help companies reduce the time, cost, and risk associated with new product introductions. Hauser *et al.* [35] observe that modern product development processes are designed to take advantage of common components and discuss an example in which component reuse resulted in a dramatic increase in profits. As one of their research challenges, the authors note that marketing incentives are disconnected from product development incentives and stress the need to develop practical models for setting and adjusting priorities for innovation. Our mathematical optimization model, by capturing market effects as well as operational characteristics, aims to be a holistic approach for making decisions in this domain.

III. Basic Line Extension Model

Section III-A introduces the basic model, referred to as Problem RS, developed in [20]. In Section III-B, we develop a simple algorithm for Problem RS that offers an attractive performance guarantee. Then, based on analytical results, we revise this algorithm in Section III-C to obtain a more sophisticated and robust heuristic.

A. Problem Formulation

Typically, a firm conducts a preliminary evaluation to identify a set of potential line extensions. We let N denote the cardinality of this set. Then, the firm's task is to select a subset of this set for introduction into the market. We now define the parameters and variables of the problem.

Parameters

N Number of (potential) line extensions.
M Number of components.
Q_k Estimated life-cycle sales for line extension k.
ΔR_k Revenue from introducing line extension k.
d_k Product-specific development cost for line extension k.
g_k Total life-cycle support cost for line extension k.
o_k Per unit labor cost for line extension k.
d_c^C Component-specific development cost for component c.
m_c Per unit material cost for component c.
$\bar{Q}_c$ Sum of the demand of all the line extensions containing component c.
$l_c^h \bar{Q}_c$ The maximum labor cost of component c.
E_c Critical production volume for component c.
l_c^h Higher per unit labor cost for component c (when its cumulative production volume is less than E_c).
l_c^l Lower per unit labor cost for component c (when its cumulative production volume is greater than E_c).
S_c^C Set of line extensions that require component c.
S_k Set of components that are required to assemble line extension k.
ρ Density of the product–component graph. That is, $\rho = (\sum_{k=1}^{N} S_k)/(N \times M)$.

Variables

X_k Line extension indicator: 1 if extension k is introduced; 0 otherwise.
γ_c Component indicator: 1 if component c is introduced; 0 otherwise.
W_c Critical production volume indicator: 1 if the cumulative production volume for component c exceeds its critical production volume E_c; 0 otherwise.
V_c^h Production volume of component c at the higher labor cost.
V_c^l Production volume of component c at the lower labor cost.

Component sharing affects the cost of line extensions at the product level, but this effect is not directly traceable to individual components. Recognizing that the revenue (respectively, cost) impact of line extensions is best evaluated at the product (respectively, component) level, Ramdas and Sawhney [20] first develop a source-of-volume model to estimate the sales of line extensions under cannibalization and then formulate a mixed-integer programming model using these estimates to identify a subset of line extensions that maximizes the profit. The demand Q_k for each line extension k—estimated using a source-of-volume model—is partitioned into three types.

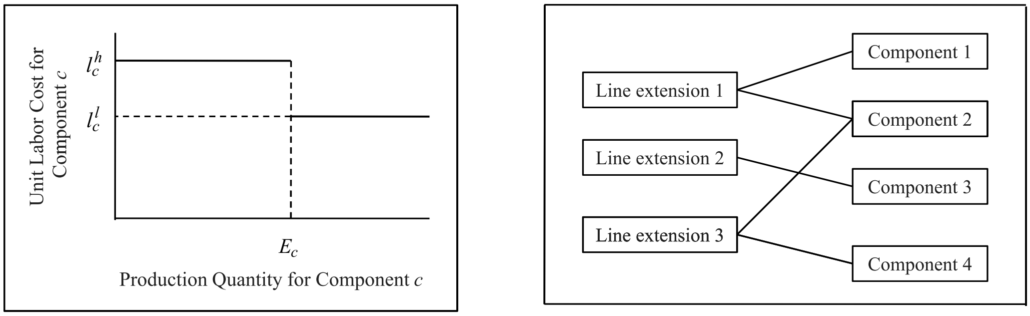
1) *Demand expansion* is sales from new consumers.
2) *Competitive draw* is sales drawn from the competition.
2) *Cannibalization* is sales drawn from consumers who would otherwise purchase the firm's existing products. We now briefly state the estimation procedure for cannibalization.

Let B_f (respectively, B_c) denote the set of exiting products offered by the firm (respectively, its competitors). Thus, $B = B_f \cup B_c$ is the set of existing products in the market. For a product $j \in B$, let $P(j|B)$ (respectively, $P(j|\{B,k\})$) be the share of preference for product j in the set B (respectively, $\{B,k\}$) measured before (respectively, after) introducing line extension k; these can be estimated by carefully designed surveys. Let Q_j denote the existing sales of product j and Q_k^C denote the estimated sales of line extension k due to cannibalization. Then, we have

$$Q_k^C = \sum_{j \in B_f} \frac{P(j|B) - P(j|\{B,k\})}{P(j|\{B,k\})} Q_j.$$

The estimated sales of line extension k due to demand expansion (Q_k^D) and competitive draw (Q_k^M) are obtained in a similar manner. Finally, we have $Q_k = Q_k^C + Q_k^D + Q_k^M$. Consequently, the model assumes that Q_k and the revenue $\Delta R_k = p_k Q_k$ are known for each line extension k, where p_k is the exogenous price for line extension k. For the detailed demand-estimation procedures, we refer the reader to [20].

The cumulative labor cost is typically a concave function of the total production volume, in the sense that the marginal cost usually decreases with the production volume due to workers' improved learning. In [20], this concavity is approximated as a two-piece linear function: the per unit labor cost for a component



(a) (b)

Fig. 1. (a) Two-piece linear labor cost for component c. (b) Illustrative example of the product–component bipartite graph.

c drops from l_c^h to l_c^l when the total production volume for this component exceeds a known critical volume E_c [see Fig. 1(a)].

A bipartite relationship between line extensions and their components (i.e., the components of each line extension) is assumed to be known. Fig. 1(b) illustrates an example of the product–component relationship graph. For simplicity, we assume that each line extension uses at most one unit of each component.[1] The product-specific costs consist of the labor cost (o_k, per unit of production), the development cost (d_k), and the support cost (g_k), which is the indirect cost (e.g., maintenance, engineering) incurred during a product's life cycle. The component-specific costs consist of the material cost (m_c, per unit of production), the development cost (d_c^C), the higher-level labor cost (l_c^h, per unit of production), and the lower-level labor cost (l_c^l, per unit of production).

Our objective is to select a subset of line extensions to maximize the total profit, which equals the total revenue minus the total cost incurred in introducing the chosen line extensions. A mixed-integer programming model of the problem is as follows:

$$\text{Problem RS}: \text{ Max} \sum_{k=1}^{N} \{\Delta R_k - d_k - g_k - o_k Q_k\} X_k - \sum_{c=1}^{M} d_c^C \gamma_c - \sum_{c=1}^{M} m_c (V_c^h + V_c^l) - \sum_{c=1}^{M} (l_c^h V_c^h + l_c^l V_c^l)$$

$$\text{Subject to: } \gamma_c \geq X_k \quad \forall k, \forall c \in S_k, \tag{1}$$

$$V_c^h + V_c^l = \sum_{k \in S_c^C} Q_k X_k \quad \forall c, \tag{2}$$

$$E_c W_c \leq V_c^h \leq E_c \quad \forall c, \tag{3}$$

$$V_c^l \leq \bar{Q}_c W_c \quad \forall c, \tag{4}$$

$$V_{c^h},\ V_c^l \geq 0 \quad \forall c, \tag{5}$$

$$X_k, \gamma_c, W_c \in \{0, 1\} \quad \forall k, c.$$

[1]The model can be easily generalized to address the situation when line extensions use multiple units of some components. Let α_{ck} denote the number of units of component c in line extension k. Then, Constraint (2) below is modified as follows: $V_c^h + V_c^l = \sum_{k \in S_c^C} (\alpha_{ck} Q_k) X_k$.

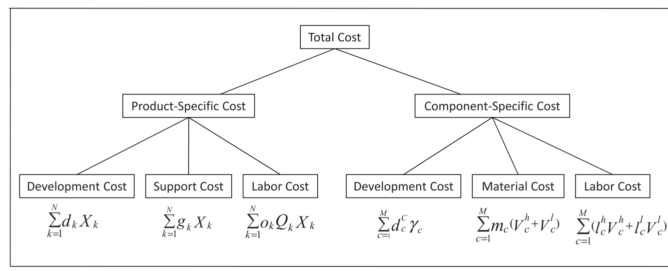


Fig. 2. Cost structure for Problem RS.

The total cost consists of product-specific costs (including the development cost, support cost, and labor cost) and component-specific costs (including the development cost, material cost, and labor cost); Fig. 2 illustrates the corresponding terms. As mentioned earlier, the component-specific labor costs are approximated by piecewise linear functions, with each consisting of two linear regions. Constraint (1) imposes that if a line extension k is selected, then all the components required for this extension must also be selected. Constraints (2) through (5) specify the production volumes V_c^h and V_c^l of a component c at the higher labor cost (l_c^h) and the lower labor cost (l_c^l), respectively.

B. Approximation Algorithm for Problem RS

Despite our serious attempts to resolve the complexity of Problem RS, this question remains open. However, we establish the hardness of both the budget-constrained and cardinality-constrained variants in Section IV and Appendix B, respectively. As can be seen later from our computational experience in Section V, industry-size instances of Problem RS as well as the two variants are challenging to solve to optimality or near-optimality. We, therefore, start by developing an effective algorithm for Problem RS. Later, in Section IV and Appendix B, we extend this algorithm for the two variants. Our algorithm uses a relaxation of Problem RS that exploits the concept of total unimodularity of the constraint matrix of a linear integer program.

If we relax the assumption of a two-piece labor cost (l_c^h and l_c^l) for component c and, instead, use a single per-unit labor cost, say l_c, then Problem RS becomes

$$\begin{aligned}\text{Problem P}_R:\ \text{Max} \sum_{k=1}^{N} \{ & \Delta R_k - d_k - g_k - o_k Q_k \\ & - \sum_{c \in S_k} (m_c + l_c)) Q_k \} X_k \\ & - \sum_{c=1}^{M} d_c^C \gamma_c. \\ \text{Subject to: } & \gamma_c \geq X_k \quad \forall k, \forall c \in S_k, \\ & X_k, \gamma_c \in \{0,1\}, \quad \forall k, c. \end{aligned} \tag{6}$$

The following result comments on the solvability of Problem P_R. The proofs of all technical results are in Appendix C.

Theorem 1: Problem P_R is polynomially solvable.

Now consider the following algorithm, which solves an instance of Problem P_R.

Algorithm Cost-Relax:

Given an arbitrary instance I of Problem RS, construct a corresponding instance, say I_R^h, of Problem P_R by using $l_c = l_c^h$, the higher per-unit labor cost, for each component c. Let S_h be the optimal set of line extensions for Instance I_R^h. We use S_h as an approximation for the optimal set of line extensions for Instance I.

To analyze the performance of this algorithm, we introduce three parametric constants.

1) Let α, $0 < \alpha < 1$, be the maximum possible percentage discount on component-specific labor costs. That is, $\alpha = \max_c \{ \frac{l_c^h - l_c^l}{l_c^h} \}$.
2) For an instance I of Problem RS, let I_R^l be an corresponding instance of Problem P_R by using $l_c = l_c^l$, the lower per-unit labor cost, for each component c. Let φ_l and C_l denote the revenue and the total cost, respectively, corresponding to the optimal set of line extensions for Instance I_R^l, and let $\beta > 0$ be such that $\varphi_l = (1+\beta) C_l$. If α is low, then Instance I_R^l can be considered as a reasonable approximation of Instance I. Thus, in this case, β can be loosely interpreted as the *percentage markup* (i.e., the ratio of profit to cost) for Instance I.
3) Let μ be the maximum ratio of the component-specific labor cost (computed via the lower unit labor cost for each component) to the total cost (except the component-specific development cost) of introducing a line extension. That is,

$$\mu = \max_k \left\{ \frac{\sum_{c \in S_k} l_c^l Q_k}{o_k Q_k + d_k + g_k + \sum_{c \in S_k} m_c Q_k + \sum_{c \in S_k} l_c^l Q_k} \right\}.$$

Let f^* denote the optimal profit of Instance I and f_h be the profit corresponding to set S_h for Instance I using Algorithm Cost-Relax. The result below establishes a nontrivial guarantee for Algorithm Cost-Relax, when $\frac{\alpha\mu}{1-\alpha} < \beta$. Intuitively, this result shows that solving Problem RS by ignoring the discounts on labor costs is guaranteed to result in a good solution if the percentage discounts on the labor cost for the components are small (e.g., ranging between 0% and 20%), the ratio of the component-specific labor cost to the total cost of introducing a line extension is small (e.g., 10–20%), and the percentage markup is high (e.g., 20–30%). For instance, when these three values are 15%, 20%, and 20%, respectively, the profit corresponding to the solution of Algorithm Cost-Relax is guaranteed to be at least 82% of the optimal profit.

Theorem 2: Algorithm Cost-Relax is an $\frac{\alpha\mu}{(1-\alpha)\beta}$-approximation. That is, $\frac{f_h}{f^*} \geq 1 - \frac{\alpha\mu}{(1-\alpha)\beta}$.

For better performance in practice, our next task is to improve Algorithm Cost-Relax.

C. Two Nesting Results: An Improved Heuristic for Problem RS

While Theorem 2 establishes a nontrivial performance bound for Algorithm Cost-Relax, the solution can be further improved based on two structural properties of Problem RS. We first develop the required notation and illustrate the basic idea.

Let $\mathcal{K}$ denote the set of all line extensions. An arbitrary feasible solution of Problem RS corresponds to a subset of chosen line extensions $S \subseteq \mathcal{K}$. For an arbitrary positive integer n and an arbitrary instance I of Problem RS with the two-piece labor cost (i.e., l_c^h and $l_c^l; l_c^l < l_c^h$) for each component c, let $l_c^i, i = 1, 2, \ldots, n$, be such that $l_c^h > l_c^1 > l_c^2 > \cdots > l_c^n > l_c^l$. Let I_R^h (respectively, I_R^i, $i = 1, 2, \ldots, n$; I_R^l) be the corresponding instance of Problem P_R with the linear labor cost $l_c = l_c^h$ (respectively, l_c^i, $i = 1, 2, \ldots, n$; l_c^l), instead of the original two-piece labor cost. Let S^* (respectively, S_h; S_i, $i = 1, 2, \ldots, n$; S_l) be the optimal set of line extensions for instance I (respectively, I_R^h; I_R^i, $i = 1, 2, \ldots, n$; I_R^l). Note that instances I_R^h, I_R^1, $\mathrm{I}_R^2, \ldots, \mathrm{I}_R^n$, and I_R^l, are polynomially solvable (see Theorem 1). Theorem 3 shows that the optimal set S^* of Problem RS is a *subset* of S_l and a *superset* of S_h. Then, Theorem 4 establishes a nesting property: $S_h \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n \subseteq S_l$ [see Fig. 3(a)]. These results provide two insights: first, for a reasonably large n, the best of the sets $S_h, S_1, S_2, \ldots, S_n, S_l$ (i.e., the one that offers the highest profit for Problem RS) can be expected to be a good approximation for S^*. Second, the quality of this approximation can be expected to improve as n increases.

Theorem 3: $S_h \subseteq S^* \subseteq S_l$.

The proof of the following result is similar to that of Theorem 3.

Theorem 4: $S_h \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n \subseteq S_l$.

Based on Theorems 3 and 4, we can now improve Algorithm Cost-Relax as follows: for a positive integer $L \geq 1$, we generate $L+1$ problems $\mathrm{P}^i, i = 0, 1, \ldots, L$. Problem P^i uses the linear labor cost $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$, which is a convex combination of the higher labor cost l_c^h and the lower labor cost l_c^l. Note that Problems $\mathrm{P}^i, i = 0, 1, \ldots, L$, are polynomially solvable (see Theorem 1). The performance of the heuristic improves with an increase in L; however, the computational time required increases as well. For $i = 0, 1, \ldots, L$, an optimal solution of Problem P^i is further enhanced using a procedure IMPROVE-RS, which exploits the sharing of components between the line

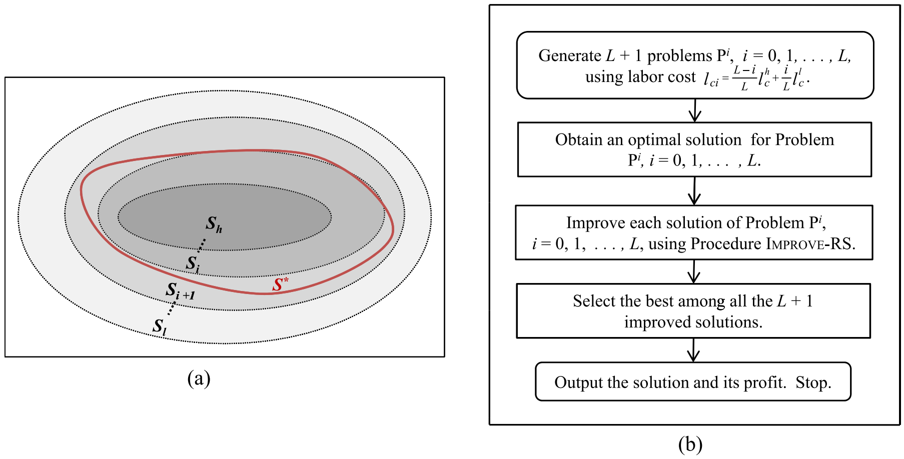


Fig. 3. (a) Relationship between sets $S_h, S_1, S_2, \ldots, S_n, S_l$ and the optimal set S^*. (b) Flowchart of heuristic H_{RS}.

extensions. We, therefore, have $L+1$ improved solutions and select the best as the final solution of our heuristic. A flowchart of the heuristic is presented in Fig. 3(b).

We now formally describe the heuristic. Let $Z_k = 1$, if extension k is chosen; 0 otherwise.

Heuristic H_{RS}

Input: An instance of Problem RS with N products and M components. A positive integer L.

Output: A solution $Z_k, k = 1, 2, \ldots, N$, of Problem RS and the corresponding profit f.

1) Step 0 (initialization): $Z_k = 0, k = 1, 2, \ldots, N$; $f = 0$, and $i = 0$.
2) Step 1: While $i \le L$, perform the following:
   - Set $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$. Generate an instance of Problem P_R by using the linear labor cost l_{ci} and obtain its optimal solution by solving the corresponding linear programming (LP) relaxation. Let this solution be $X_k^i, k = 1, 2, \ldots, N$.
   - Use procedure IMPROVE-RS to obtain an enhanced solution $\bar{X}_k^i, k = 1, 2, \ldots, N$, and its profit $\bar{f}_i$. If $\bar{f}_i > f$, set $f = \bar{f}_i$ and $Z_k = \bar{X}_k^i, k = 1, 2, \ldots, N$.
   - Set $i = i + 1$.
3) Step 2: Output the solution $Z_k, k = 1, 2, \ldots, N$, and its profit f.

Procedure IMPROVE-RS

Input: A feasible solution $X_k, k = 1, 2, \ldots, N$, of Problem RS.

Output: An improved solution $\bar{X}_k, k = 1, 2, \ldots, N$, of Problem RS and the corresponding profit $\bar{f}$.

1) Step 0: Set $\bar{X}_k = X_k, k = 1, 2, \ldots, N$ and $I = 0$.
2) Step 1: While $I = 0$, perform the following:
   - Set $I = 1$, $c = 0$, and $k = 0$.
   - While $c \le M$, perform the following:
     i) If $\sum_{k \in S_c^C} \bar{X}_k > 0$, then set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{k \in S_c^C} Q_k \bar{X}_k\}$.
     ii) $c = c + 1$.
   - While $k \le N$, perform the following:
     i) If $\bar{X}_k = 0$, then recompute its profit $\triangle f_k$ using the redefined d_c^C and E_c. If $\triangle f_k > 0$, then set $\bar{X}_k = 1$ and $I = 0$.
     ii) $k = k + 1$.
3) Step 2: Compute the profit $\bar{f}$ of the new solution $\bar{X}_k, k = 1, 2, \ldots, N$.

Next, we describe the idea behind the procedure IMPROVE-RS used in the above heuristic. Recall that $\mathcal{K}$ is the set of all line extensions. For an arbitrary set of chosen line extensions $S \subseteq \mathcal{K}$, let $\mathcal{C}(S)$ be the set of components used by the extensions in S. Now, consider an extension $k \in \mathcal{K} \backslash S$. An interesting question is whether or not the solution improves if extension k is also introduced. Let

$$\begin{aligned} \triangle f_k = \Delta R_k - o_k Q_k - d_k - g_k \\ - \sum_{c \in S_k} d_c^C - \sum_{c \in S_k} m_c Q_k - \sum_{c \in S_k} l_c^h \min\{Q_k, E_c\} \\ - \sum_{c \in S_k} l_c^l \max\{0, Q_k - E_c\}. \end{aligned}$$

Our next result answers the question above. The proof is straightforward. Therefore, for brevity, we avoid stating it explicitly.

Theorem 5: Let S be a set of chosen extensions of Problem RS with profit f. For each $c \in \mathcal{C}(S)$, redefine $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{k \in S_c^C \cap S} Q_k\}$. For a line extension $k \in \mathcal{K} \backslash S$, the profit corresponding to the set $S \cup \{k\}$ is $f + \triangle f_k$. Thus, expanding S by including extension k improves the profit if $\triangle f_k > 0$.

This result immediately suggests an iterative improvement procedure. Starting with a set S, we can choose a line extension $k \in \mathcal{K} \backslash S$ with $\triangle f_k > 0$ (if one exists) to obtain an improved set $S \cup \{k\}$. This process can be repeated until no improvement is possible.

Remark: The basic idea of our heuristic can be easily extended to the more general case when a piecewise linear labor cost has $n \ge 3$ segments. Suppose the unit labor cost for a component c is specified by the set $\{l_c^1, l_c^2, \ldots, l_c^n\}$ (instead of the simpler set $\{l_c^h, l_c^l\}$ of the current model), where $l_c^1 \ge l_c^2 \ge \cdots \ge l_c^n$ are the n values corresponding to their respective ranges of the production quantity. Then, the first step of Heuristic H_{RS} can be modified as follows. For a positive integer $L \ge 1$, we generate

$(n-1)L+1$ problems: Problem $\mathrm{P}_k^i, i=0,1,\ldots,L-1$; $k=1,2,\ldots,n-1$ and Problem P_{n-1}^L. Problem P_k^i uses the linear labor cost $l_{ck}^i = \frac{L-i}{L} l_c^k + \frac{i}{L} l_c^{k+1}$, which is a convex combination of the labor costs l_c^k and l_c^{k+1}, and Problem P_{n-1}^L uses the linear labor cost l_c^n. Again, Problems P_k^i $(i=0,1,\ldots,L-1;$ $k=1,2,\ldots,n-1)$ and P_{n-1}^L are polynomially solvable (see Theorem 1).

In the context of new product development, firms are typically constrained either by the *budget* available to introduce the line extensions or the *number* of line extensions to be chosen. In the next section, we analyze the budget-constrained variant (P_b) of Problem RS. The cardinality-constrained variant (P_c) is introduced at the end of Section IV-A, but is discussed in detail in Appendix B.

IV. Line Extensions Under a Budget Constraint: Problem P_b

Let B denote the imposed budget on the total cost. Our aim in Problem P_b is to select a subset of line extensions from a desirable set to maximize the total profit, under the constraint that the total cost incurred is at most B. We resolve the complexity status of Problem P_b and then develop an effective and robust heuristic.

In addition to the constraints of Problem RS (see Section III-A), the formulation of Problem P_b includes the following budget constraint:

$$\sum_{k=1}^{N} X_k\{o_k Q_k + d_k + g_k\} + \sum_{c=1}^{M} d_c^C \gamma_c + \sum_{c=1}^{M}\{(m_c + l_c^h)V_c^h + (m_c + l_c^l)V_c^l\} \leq B.$$

The following result resolves the complexity of Problem P_b.

Theorem 6: The decision problem corresponding to P_b is NP-complete, even when the number of components is one.

A. *Heuristic for Problem P_b*

The idea behind our heuristic H_b for Problem P_b is similar to that for Heuristic H_{RS} (see Section III-C). The basic property exploited is as follows: if, instead of a two-piece unit labor cost for each component c (i.e., the lower labor cost l_c^l and the higher labor cost l_c^h), we only have a one-piece unit labor cost, then the optimization problem can be solved by an LP (see Theorem 1). Thus, the idea is to approximate the two-piece cost function by a linear cost function. The heuristic generates several linear cost functions by considering convex combinations of l_c^l and l_c^h and solves several LPs.

In Step 1 of the heuristic, we first generate $L+1$ problems $\mathrm{P}^i, i=0,1,\ldots,L$, by using the following convex combinations of the two costs: $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$. If the cost incurred in an optimal solution of Problem P^i is less than B, then we use Procedure Improve-P_b to add more line extensions and increase the total profit; otherwise, if the cost is greater than B, then we use Procedure Reduce-P_b to remove some line extensions to satisfy the budget constraint. Let A_1 be the best among these $L+1$ solutions. Then, in Step 2 of the heuristic, starting with only one line extension $k, k=1,2,\ldots,N$, as the initial solution, we use Procedure Improve-P_b to enhance it. Let A_2 denote the best among these N solutions. We then select the better of A_1 and A_2 as the solution of the heuristic. We now formally describe the heuristic.

<u>Heuristic H_b</u>

Input: An instance of Problem P_b with N products and M components. A positive integer L.

Output: A solution $Z_k, k=1,2,\ldots,N$, and its corresponding profit f.

1) Step 0 (initialization): $Z_k=0, k=1,2,\ldots,N$; $f=0$, and $i=0$.
2) Step 1: While $i \leq L$, perform the following:
   - Set $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$. Generate an instance of Problem P_R by using the linear labor cost l_{ci} and obtain its optimal solution by solving the corresponding LP relaxation. Let this solution be $X_k^i, k=1,2,\ldots,N$.
   - Let B^i be the total cost corresponding to $X_k^i, k=1,2,\ldots,N$. If $B^i < B$, use Procedure Improve-P_b to enhance the solution. Otherwise, if $B^i > B$, use Procedure Reduce-P_b (Appendix A) to obtain a feasible solution. Let the improved solution be $\bar{X}_k^i, k=1,2,\ldots,N$, with corresponding profit $\bar{f}_i$. If $\bar{f}_i > f$, set $f=\bar{f}_i$ and $Z_k = \bar{X}_k^i, k=1,2,\ldots,N$.
   - Set $i=i+1$.
3) Step 2: Set $i=0$. While $i \leq N$, perform the following:
   - Set $X_i=1$ and $X_j=0, j\neq i$.
   - Use Procedure Improve-P_b (Appendix A) to enhance the solution. Let the improved solution be $\bar{X}_k^i$, $k=1,2,\ldots,N$, and its profit be $\bar{f}_i$. If $\bar{f}_i > f$, set $f=\bar{f}_i$ and $Z_k=\bar{X}_k^i, k=1,2,\ldots,N$.
   - $i=i+1$.
4) Step 3: Output the solution $Z_k, k=1,2,\ldots,N$, and its profit f.

We summarize the idea behind procedures Improve-P_b and Reduce-P_b . The formal descriptions of these two procedures are provided in Appendix A. For a set S of Problem P_b, let $\mathcal{C}(S)$ be the components used by S and let $B(S)$ be the cost of introducing the line extensions in S. For $c\in\mathcal{C}(S)$, we update the development cost to 0 and the critical production volume to $\max\{0, E_c - \sum_{k\in S_c^C, k\in S} Q_k\}$. If $B(S) < B$, we add line extensions to S using the following two rules and then select the better of the resulting two solutions.

1) Add a feasible line extension with the largest positive incremental profit.
2) Add a feasible line extension with the largest ratio of the incremental profit to the incremental cost.

If $B(S) > B$, we again use these two rules to remove line extensions until the budget constraint is satisfied, and then select the better solution.

We end this section by briefly introducing the cardinality-constrained variant P_c. Let $U\in Z_+$ denote an upper bound on the number of line extensions to be introduced. Thus, our aim in Problem P_c is to select a subset of line extensions from a desirable set to maximize the total profit, under the constraint that the total number of selected line extensions is at most U.

Except for the cardinality constraint $\sum_{k=1}^{N} X_k \leq U$, the other constraints are the same as in Section III-A. Our approach for Problem P_c is similar to that for Problem P_b, and is discussed in Appendix B.

In the next section, we examine the performance of our heuristics on a comprehensive test bed of instances.

V. Computational Experience

Recall that Problem RS is a special case of Problem P_b (respectively, P_c), corresponding to the case when the upper bound on the budget (respectively, cardinality) is large enough. Therefore, we focus only on Problems P_b and P_c. First, to justify the need for a heuristic, we develop a set of computationally hard instances of Problems P_b and P_c. CPLEX (version 11.1.1, running on a 2.8-GHz Intel XEON with 1-GB RAM) is unable to solve several of these instances to optimality or near-optimality within a reasonable time limit (2 h of CPU time). Then, on a comprehensive test bed of realistic instances, we compare our heuristics to popular rule-of-thumb approaches used in the industry. We show that our heuristics quickly deliver near-optimal solutions that are significantly superior to those obtained from these common approaches.

A. *Need for an Effective Heuristic*

Recall from Section III-A the following parameters: 1) ρ ($0 < \rho < 1$) is the density of the given product–component bipartite graph, 2) $\bar{Q}_c$ is the sum of the demands of all the line extensions containing component c, 3) E_c is the critical production volume for component c, and 4) $l_c^h \bar{Q}_c$ as the maximum labor cost of component c.

We now describe the test bed used in our computations. The values of the parameters ρ, D, e, λ, t, η, ϕ, and δ, in the description below are specified later for individual experiments. Each line extension includes a randomly chosen set of ρM components. The demand Q_k for line extension k is randomly chosen from U[4000, 6000], where U$[a, b]$ is the uniform distribution over the interval $[a, b]$. For each component c, the higher unit labor cost l_c^h is randomly chosen from U[0, 20] and the lower unit labor cost is set as $l_c^l = D l_c^h$ ($0 < D < 1$). The critical production volume for component c is set as $E_c = e\bar{Q}_c$ ($0 < e < 1$) and the development cost of the component c is set as $d_c^C = \lambda l_c^h \bar{Q}_c$ ($0 < \lambda < 1$). For simplicity, we set the unit material cost $m_c = 0$ for each component c. For each line extension k, the fixed cost (i.e., the sum of the development and support cost; see Section III-A) is set equal to t times ($0 < t < 1$) the total development costs of all the components it includes. The unit product-specific labor cost o_k is a random number chosen from U[0, 100]. Let C_k^h (respectively, C_k^l) be the total cost of introducing line extension k (i.e., the sum of product-specific labor costs, development costs and supporting costs, and labor costs and development costs for all the components included in this line extension) obtained by using the higher (respectively, lower) component-specific labor costs. Then, the revenue for extension k is $R_k = C_k^l + \eta(C_k^h - C_k^l)$, where η is a random number whose generation will be specified for individual experiments. Finally, for $0 < \phi < 1$ and $0 < \delta < 1$, 1) the available budget (see Section IV) is $B = \phi R$, where R is the total revenue from introducing all line extensions and 2) the upper bound on the cardinality of the line extensions (see Section IV-A and Appendix B. is $U = \delta N$. Thus, ϕ represents the maximum allowable ratio of the cost incurred in introducing the selected line extensions to the total revenue from introducing all possible line extensions and δ represents the maximum allowable ratio of the number of chosen extensions to the total number of line extensions.

TABLE I
Average Percentage Gaps of the Solutions from CPLEX for Hard Instances of Problem P_c and P_b

(a) Problem P_c

$\frac{M}{N}$ ↓	$N = 30$	$N = 50$	$N = 70$	$N = 100$
0.5	0.0%	0.0%	72.9%	92.7%
1.0	0.0%	97.1%	98.2%	99.3%
1.5	65.6%	97.6%	100.0%	100.0%
2.0	97.7%	98.2%	100.0%	100.0%

(b) Problem P_b

$\frac{M}{N}$ ↓	$N = 30$	$N = 50$	$N = 70$	$N = 90$
0.5	0.0%	0.0%	93.6%	95.1%
1.0	11.9%	95.7%	98.4%	99.7%
1.5	95.4%	100.0%	100.0%	100.0%
2.0	99.2%	100.0%	100.0%	100.0%

To generate a class of hard instances for Problems P_b and P_c, we set $\rho = 0.5, D = 0.9, e = 0.5, \lambda = 0, t = 0, \eta \in \mathrm{U}[0.7, 1.0]$, $\phi = 0.5$, and $\delta = 0.5$. For an instance that is not solved to optimality by CPLEX within 2 h of CPU time, the percentage gap is defined as shown, at the bottom of the page.

Each entry in Table I(a) and (b) shows the average percentage gap of two instances. A similar class of computationally challenging instances can be constructed for Problem RS as well. It is important to note that, for an assembled product, there are typically hundreds of possible line extensions that are evaluated each year; see, for example, the discussion in [20] of the products of Titan Industries Limited, an international manufacturer of wrist watches. Thus, there is a strong need for an effective and efficient heuristic.

B. *Performance of Heuristics H_c and H_b*

For Problem P_c, Ramdas and Sawhney [20] describe two common heuristics, $\mathrm{H}^{\mathrm{REV}}$ and $\mathrm{H}^{\mathrm{ROI}}$, used in the industry. We consider a natural extension of these heuristics for Problem P_b. We describe them briefly below:

$$\%\text{Gap} = \frac{(\text{Best Available Upper Bound} - \text{Best Feasible Solution Found}) \times 100\%}{\text{Best Available Upper Bound}}.$$

TABLE II
PARAMETER SETTINGS TO EVALUATE THE PERFORMANCE OF THE HEURISTICS

(a) Small Size Problems

Parameters	Values
(N, M)	(10,10), (10,30), (30,15), (30,30), (30,60)
ρ	0.2, 0.5, 0.8
D	0.5, 0.8
e	0.2, 0.5, 0.8
(λ, t)	(0, 0), (0.3, 0), (0.3, 0.5)
δ	0.2, 0.5, 0.8
ϕ	0.2, 0.5, 0.8
Q_k	U[4000, 6000]
l_c^h	U[0, 20]
o_k	U[0, 100]
η	U[0.5,1.5]

(b) Large Size Problems

Parameters	Values
(N, M)	(100,200)
ρ	0.2, 0.5, 0.8
D	0.5, 0.7, 0.9
e	0.2, 0.5, 0.8
(λ, t)	(0.1, 0.1)
δ	0.5
ϕ	0.5
Q_k	U[4000, 6000]
l_c^h	U[0, 20]
o_k	U[0, 100]
η	U[0.5, 1.5]

TABLE III
PERFORMANCE OF THE HEURISTICS FOR SMALL-SIZE PROBLEMS

Metric	Problem	H^{REV}	H^{ROI}	H^{RR}	H_c	H_b
Average Gap	Problem P_c	11.48%	8.69%	4.21%	0.04%	–
	Problem P_b	8.70%	9.18%	3.85%	–	0.17%
Maximum Gap	Problem P_c	100.00%	95.90%	95.90%	6.20%	–
	Problem P_b	100.00%	100.00%	76.30%	–	9.80%

TABLE IV
(a) PERCENTAGE OF INSTANCES THAT ARE SOLVED TO OPTIMALITY FOR SMALL-SIZE PROBLEMS. (b) PERCENTAGE IMPROVEMENT O THE PROFIT PROVIDED BY OUR HEURISTIC OVER THOSE OFFERED BY HEURISTIC H^{REV}, H^{ROI}, AND H^{RR} FOR LARGE-SIZE PROBLEMS

(a)

Problem	H^{RR}	H_c	H_b
Problem P_c	37.00%	98.00%	–
Problem P_b	65.80%	–	90.60%

(b)

Problem	H^{REV}	H^{ROI}	H^{RR}
Problem P_c	15.97%	8.12%	7.85%
Problem P_b	10.34%	9.92%	4.82%

HEURISTIC H^{REV}

Revenue-based selection: Rank the line extensions based on revenue and then select a set of extensions based on the rankings, subject to the condition that the total number (respectively, cost) of selected extensions is less than or equal to U (respectively, B).

HEURISTIC H^{ROI}

ROI-based selection: Rank line extensions based on ROI and then select a set of extensions based on the rankings, subject to the condition that the total number (respectively, cost) of selected extensions is less than or equal to U (respectively, B). Here, ROI is defined as the ratio of the revenue less the labor and material costs to the total development and support costs.

We denote the heuristic that chooses the better of H^{REV} and H^{ROI} by H^{RR}. As can be seen from Table I(a) and (b), the relatively smaller instances are solvable to optimality within a reasonable amount of time. For such instances, we compare the average gap between the profit corresponding to the heuristic solution and the optimal profit. For larger instances, we evaluate the percentage improvement in the profit provided by our heuristics over that offered by Heuristics H^{REV}, H^{ROI}, and H^{RR}. Table II(a) [respectively, Table II(b)] indicates the values of parameters for small (respectively, large) size problems. For each combination of parameters for the small (respectively, large) size problems, we generate one instance (respectively, five instances) of Problem P_c and one instance (respectively, five instances) of P_b, for a total of 1620 (respectively, 270) instances.

Table III shows the average and maximum gaps—over the 810 instances each of Problems P_c and P_b—between the profit corresponding to the heuristic solution and the optimal profit for the small size problems. For Problem P_c, the average gap of the solutions from Heuristic H^{REV} (respectively, H^{ROI}, H^{RR}) is 11.48% (respectively, 8.69%, 4.21%). In comparison, the average optimality gap of the solutions from Heuristic H_c is 0.04%. Similarly, the maximum gap of the solutions from Heuristic H^{REV} (respectively, H^{ROI}, H^{RR}) is 100.00% (respectively, 95.90%, 95.90%), while that of the solutions from Heuristic H_c is 6.20%. Moreover, Heuristic H_c provided the solution of each instance within 5 s of CPU time. The results of Heuristic H_b for Problem P_b are similar.

Table IV(a) shows the percentage of instances (from the 810 instances each of P_c and P_b) that are solved to optimality. For Problem P_c, the solution of Heuristic H^{RR} was optimal for 37% of the instances, while Heuristic H_c offered an optimal solution for 98% of the instances. Again, the results are similar for Problem P_b.

Next, we compare the heuristics for large-size problems. As shown in Table IV(b), for Problem P_c, the percentage increase from the profit offered by Heuristic H^{REV}(respectively, H^{ROI}, H^{RR}) to that provided by Heuristic H_c is 15.97% (respectively, 8.12%, 7.85%). The results for Heuristic H_b are similar. When both the density of the product–component bipartite graph and the percentage discount on the labor costs are high, the performance of our heuristics improves further. In this case, the average percentage increase from the profit offered by Heuristic H^{REV}(respectively, H^{ROI}, H^{RR}) to that provided by Heuristic H_c is 48.82% (respectively, 26.82%, 26.82%). Again, the improvement offered by Heuristic H_b is similar. Both Heuristic H^{REV} and Heuristic H^{ROI} ignore the impact of the economy of scale and consider only linear labor costs, while Heuristics H_b and H_c fully explore the advantage of lower labor costs and component

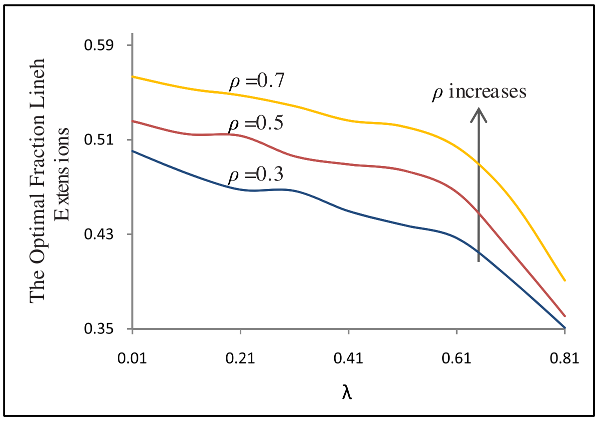


Fig. 4. Impact of 1) the ratio of the component-specific development cost to the maximum labor cost (λ) and 2) the density of the given product–component bipartite graph (ρ), on the optimal fraction of line extensions for problem P_b.

sharing. Consequently, when the differences between the higher and lower labor costs are significant, the relative advantage offered by our heuristics increases significantly. Similarly, when the density of the product–component graph is high, our heuristics exploit the opportunity to lower labor costs via the sharing of components.

C. *Three Targeted Experiments*

In this section, we discuss three targeted experiments conducted to derive several useful insights on the optimal set of line extensions.

1) Impact of Problem-Specific Characteristics on the Optimal Fraction of Line Extensions: Recall from Section V-A that λ is the ratio of the component-specific development cost to the maximum labor cost and ρ is the density of the product–component bipartite graph. We consider three values of ρ: 0.3, 0.5, and 0.7. For each value of ρ, we consider eight values of λ: 0.01, 0.11, 0.21, 0.31, 0.41, 0.51, 0.61, and 0.71. The other parameters are set as follows: $(N, M) = (30, 30)$, $(100, 200)$; $D = 0.8$; $e = 0.5$; $t = 0.1$; $\eta \in \mathrm{U}[0.5, 1.5]$; and $\phi = 0.5$. For each combination of (N, M), ρ, and λ, we generate ten instances for Problem P_b for a total of 480 instances. As in Section V-B, we use optimal solutions (obtained via CPLEX) for instances with $(N, M) = (30, 30)$ and near-optimal solutions (obtained from our heuristics) for instances with $(N, M) = (100, 200)$.

Define the *fraction of line extensions* as follows:

$$\text{Fraction of line extensions} = \frac{\text{Number of chosen line extensions}}{\text{Total number of line extensions}}.$$

We examine the impact of λ and ρ on the optimal fraction of line extensions. From Fig. 4, we observe that the optimal fraction is a concave and decreasing function of λ and an increasing function of ρ. As λ increases, the development cost of new components progressively outweighs the benefit from the line extensions that include these components; consequently, the cardinality of the optimal set of line extensions decreases. Furthermore, at a high development cost, the impact of a marginal change in the development cost is much more influential than that at a lower development cost. This behavior is consistent with the observation that products with a higher ratio of the development cost to the labor cost (e.g., automobiles, computers) typically have fewer line extensions than those for which this ratio is low (e.g., soups, beverages). When the density of the product–component graph (ρ) is high, there is a better chance (compared with the case under low density) of achieving the critical volume and exploiting the lower labor costs by introducing more extensions. Consequently, the optimal fraction of extensions is an increasing function of ρ. In our experiments, we also observe that the marginal increase in the optimal profit decreases as ρ increases, thus reflecting decreasing marginal returns from increasing component commonality among the line extensions.

2) All-Unit Versus Marginal Quantity Discount: Recall that our analysis considered a piecewise linear component-specific labor cost (motivated by workers' learning curve), but a linear (i.e., per-unit) material cost. We now consider the reverse situation: a two-piece linear material cost (motivated by scale economies) and a linear component-specific labor cost.

To help the firm select a supplier, we compare two popular discount policies for the material cost: an all-unit quantity discount and a marginal quantity discount. In both policies, a discount is offered if the purchase quantity exceeds a known threshold. In an all-unit quantity discount, a discount price is applied to the entire purchase quantity. In a marginal quantity discount (where the undiscounted price is typically lower), a discount price is applied only to the amount that exceeds the threshold. For simplicity, we only consider the cardinality-constrained version. The formulations of this problem under both the discount policies are straightforward; we, therefore, avoid specifying them explicitly here.

Let θ_c (respectively, m_c) be the undiscounted per-unit price for component c in the all-unit (respectively, marginal) quantity discount policy. We assume that the discount is offered (when the purchase quantity exceeds a known threshold) as a percentage of the undiscounted price. For comparison, we further assume that both policies share the same percentage discount. Thus, when $\theta_c \leq m_c$, it is obvious that the all-unit discount is preferable. We, therefore, only consider the case when $\theta_c > m_c$. We consider three values of the ratio θ_c/m_c: 1.05, 1.10, 1.15. Recall from Section V-A that ρ is the density of the given product–component bipartite graph. Under each value of θ_c/m_c, we consider nine values of ρ: $0.1, 0.2, \ldots, 0.9$. The other parameters are set as follows: $(N, M) = (30, 30), (100, 200)$; $D = 0.8$; $e = 0.5$; $\lambda = 0.1$; $t = 0.1$; $\eta \in \mathrm{U}[0.5, 1.5]$; and $\delta = 0.5$. For each combination of (N, M), θ_c/m_c, and ρ, we generate ten instances of Problem P_c, for a total of 540 instances.

Fig. 5 plots the behavior of the ratio (say, σ) of the optimal profit under the all-unit discount to that under the marginal discount, with respect to an increase in the density ρ and the ratio θ_c/m_c. When ρ is low, the purchase quantities of most components typically do not reach their critical volumes. Consequently, the purchases of these components are at their undiscounted prices, which are lower for the marginal discount policy. Therefore, the marginal discount policy is preferable at lower values of

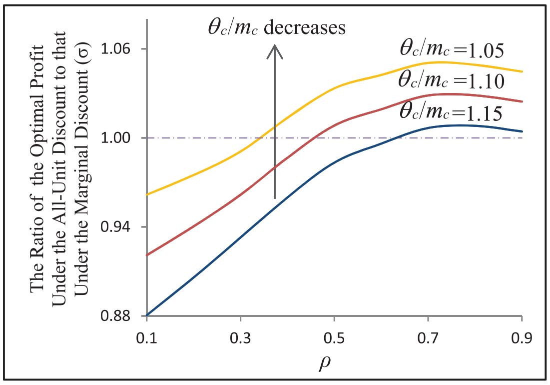


Fig. 5. Impact of the density of the given product–component bipartite graph (ρ), and the ratio of the undiscounted per-unit price in the all-unit discount to that in the marginal discount (θ_c/m_c) on the ratio of the optimal profit under the all-unit discount to that under the marginal discount for problem P_c.

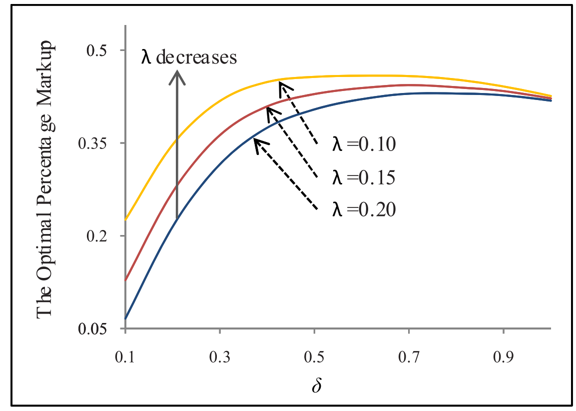


Fig. 6. Impact of the upper bound on the cardinality of line extensions (δ), and the ratio of the component-specific development cost to the maximum labor cost (λ) on the optimal percentage markup for problem P_c.

ρ. As ρ increases, the volumes start progressively exceeding their critical values and the all-unit discount policy starts becoming more desirable. When ρ is modest, the value of the ratio θ_c/m_c assumes significance: the all-unit policy is better when this ratio is close to 1. As the ratio increases, the marginal policy becomes progressively better. To summarize, the marginal (respectively, all-unit) discount is better when ρ is low (respectively, high) and ρ is modest and θ_c/m_c is large (respectively, small). Broadly, in the presence of a realistic difference between the undiscounted per-unit prices under the all-unit discount and the marginal discount policies (e.g., $1 < \theta_c/m_c \leq 1.15$), we observe that it is better to use the all-unit discount if products contain a large number of components and share significant fractions of components among themselves. Otherwise, the marginal discount is preferable. Also, as ρ increases, since purchase volumes typically exceed their critical volumes, the average (i.e., per-unit) price of the all-unit policy remains unaffected while that of the marginal policy continues to decrease. Therefore, the rate of increase of the ratio σ decreases as ρ increases. In other words, σ is a concave function of ρ.

3) Impact of Problem-Specific Characteristics on the Percentage Markup of the Optimal Solution: Let C be the total cost of introducing all line extensions (including the labor, development, and support costs for the extensions and the labor, development, and material costs for the components). For Problem P_b, let the available budget be $B = \phi C$ $(0 < \phi < 1)$. Recall from Section V-A that λ is the ratio of the component-specific development cost to the maximum labor cost, and the upper bound on the cardinality of line extensions is $U = \delta N$ $(0 < \delta < 1)$ for Problem P_c. We consider three values of λ: $0.10, 0.15$, and 0.20. For each value of λ, we consider ten values of δ (respectively, ϕ) for Problem P_c (respectively, P_b): $0.1, 0.2, \ldots, 1.0$. The other parameters are set as follows: $(N, M) = (30, 30), (100, 200)$; $\rho = 0.5$; $D = 0.8$; $e = 0.5$; $t = 0.1$; and $\eta \in$ U[0.5,1.5]. For each combination of (N, M), λ, δ (respectively, ϕ), we generate five instances each of Problem P_c and P_b. Thus, the total number of instances in this test bed is $2 \times 300 = 600$.

We examine the impact of λ and δ on the percentage markup of the optimal solution for Problem P_c. As can be seen from Fig. 6, the percentage markup is a concave function of δ and a decreasing function of λ. When the upper bound on the cardinality of extensions is relatively small ($\delta \leq 0.4$), the number of selected extensions in the optimal solution increases rapidly as δ increases. This results in a better opportunity to share development costs and exploit the lower labor costs, leading to a steady increase in the percentage markup. When δ reaches a threshold value, which is typically between 0.3 and 0.5, most of the profitable extensions have already been selected. Therefore, as δ increases further, the percentage markup stays essentially flat or even decreases. In other words, after the upper bound on the cardinality of line extensions reaches a threshold value (typically 0.3 to 0.5 times the total number of line extensions), we observe that increasing the upper bound on the cardinality does not significantly increase the percentage markup. Moreover, when λ is low, labor costs outweigh the development costs; therefore, the benefit by exploiting lower labor costs is larger (as compared to the case when λ is relatively higher). Thus, the percentage markup of the optimal solution is a decreasing function of λ. The observations for Problem P_b are similar.

VI. Insights and Future Research Directions

We summarize the insights derived from the three experiments in Sections V-C1)–V-C3). The interpretations of these insights for practicing managers are as discussed in Section I.

1) *The impact of problem-specific characteristics on the optimal fraction of line extensions:* There are two key drivers of the optimal fraction of extensions: a) the ratio of the component-specific development cost to the maximum labor cost (which corresponds to the production quantity of a component that satisfies the demand of all extensions using it); b) the density of the product–component bipartite graph. The optimal fraction of extensions is a concave and decreasing function of the ratio in a) above and an increasing function of the density of the product–component graph.

2) *Comparison of two common discount policies—all-unit (with a relatively higher undiscounted price) and marginal (with a relatively lower un-discounted price)—on the material cost incurred in producing the line extensions:* The marginal (respectively, all-unit) discount policy is better when the density of the product–component graph is low (respectively, high) or the density of the product–component graph is modest and there is a large (respectively, small) difference between the undiscounted, per-unit prices of the two policies. Moreover, the ratio of the optimal profit offered by the all-unit discount policy to that by the marginal discount policy is a concave function of the density of the product–component graph.
3) *The impact of problem-specific characteristics on the percentage markup—the ratio of the total profit to the total cost—of the optimal solution:* The percentage markup is a concave function of the upper bound on the cardinality of the line extensions (respectively, available budget) and a decreasing function of the ratio of the development cost to the labor cost.

We now conclude the paper by offering some directions for future work.

1) *Relaxing the Single-Component Assumption:* Throughout this paper, our theoretical analysis as well our numerical computations for the budget- and cardinality-constrained variants assumed that each line extension uses at most one unit of each component. As explained in Section III-A, the formulations of the two variants can be easily extended to incorporate the situation when line extensions use multiple units of some components. It may be possible to generate a much richer set of insights by analyzing these generalized formulations.
2) *Cannibalization between Line Extensions and Endogenous Pricing:* Our analysis in this paper only considered the cannibalization of demand between existing products and new line extensions. An important generalization would be to also incorporate the cannibalization*between*the line extensions. Under this addition, the tradeoff between increasing the number of line extensions (to save costs) and decreasing them (to avoid cannibalization) becomes richer. In general, the optimal number of line extensions would be less in this case (as compared to the situation without cannibalization between the extensions). Also, as a set, the significance of line extensions that share several components with each other may reduce, since such extensions are also expected to be highly similar to each other and, therefore, vulnerable to cannibalization. The problem of pricing the line extensions is also interesting and challenging. Note that, owing to the tradeoff mentioned previously, the pricing problem is relevant even under a monopolistic setting. A further enhancement could be to incorporate competition.
3) *Concave Component-Specific Labor Cost and Other Algorithmic Issues:* Our use of a piecewise linear function for the component-specific labor cost was an approximate attempt to incorporate scale economies in production. The use of a more-general, differentiable concave function can lead to an improved understanding of the optimal choice of line extensions. In some specific cases, the optimal set of extensions (and the optimal profit) exhibits an enhanced sensitivity to the parameters of our model. For instance, when labor costs dominate development costs and the critical volumes for the components are modest (e.g., about 40–60% of their maximum possible values), the optimal profit is relatively more sensitive to the labor costs. A detailed sensitivity analysis in such situations is a useful direction for future work. Finally, resolving the complexity of Problem RS is an interesting open problem.

APPENDIX A

PROCEDURES IMPROVE-P$_b$ AND REDUCE-P$_b$ (SEE SECTION IV-A)

Procedure <u>IMPROVE-P$_b$</u>

Input: A solution $X_k, k = 1, 2, \ldots, N$, of Problem P$_b$.

Output: An improved solution $\bar{X}_k, k = 1, 2, \ldots, N$, and its profit $\bar{f}$.

1) Step 0: Compute the profit f and the total cost B_1 of solution $X_k, k = 1, 2, \ldots, N$. Set $\bar{X}_{1k} = \bar{X}_{2k} = X_k, k = 1, 2, \ldots, N$, $B_2 = B_1$, and $\bar{f} = f$.
2) Step 1: While $B_1 < B$, perform the following:
   - Set $c = 0$ and $k = 0$.
   - While $c \leq M$, perform the following:
     i) If $\sum_{k \in S_c^C} \bar{X}_{1k} > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{k \in S_c^C} Q_k \bar{X_{1k}}\}$.
     ii) $c = c + 1$.
   - While $k \leq N$, perform the following:
     i) If $\bar{X}_{1k} = 0$, recompute its profit $\triangle f_k$ and cost b_k with the redefined values of d_c^C and E_c.
     ii) $k = k + 1$.
   - If there are unselected line extensions with positive profits under the budget B, add the line extension j with largest profit, set $\bar{X}_{1j} = 1$; otherwise, break.
   - $B_1 = B_1 + b_j$.
3) Step 2: Compute the profit $\bar{f}_1$ corresponding to the new solution $\bar{X}_{1k}, k = 1, 2, \ldots, N$. If $\bar{f}_1 > \bar{f}$, set $\bar{f} = \bar{f}_1$ and $\bar{X}_k = \bar{X}_{1k}, k = 1, 2, \ldots, N$.
4) Step 3: While $B_2 < B$, perform the following:
   - Set $c = 0$ and $k = 0$.
     While $c \leq M$, perform the following:
     i) If $\sum_{k \in S_c^C} \bar{X}_{2k} > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{k \in S_c^C} Q_k \bar{X_{2k}}\}$.
     ii) $c = c + 1$.
   - While $k \leq N$, perform the following:
     i) If $\bar{X}_{2k} = 0$, recompute its profit $\triangle f_k$ and cost b_k using the redefined d_c^C and E_c.
     ii) $k = k + 1$.
   - If there are unselected line extensions with positive profits under the budget B, add the line extension j with largest ratio of profit to cost, set $\bar{X}_{2j} = 1$; otherwise, break.
   - $B_2 = B_2 + b_j$.

5) Step 4: Compute the profit $\bar{f}_2$ for the new solution $\bar{X}_{2k}, k = 1, 2, \ldots, N$. If $\bar{f}_2 > \bar{f}$, set $\bar{f} = \bar{f}_2$, and $\bar{X}_k = \bar{X}_{2k}, k = 1, 2, \ldots, N$.

Procedure REDUCE-P$_b$

Input: An infeasible solution $X_k, k = 1, 2, \ldots, N$ of Problem P$_b$.

Output: A feasible solution $\bar{X}_k, k = 1, 2, \ldots, N$ and its profit $\bar{f}$.

1) Step 0: Compute the profit f and the total cost B_1 of solution $X_k, k = 1, 2, \ldots, N$. Set $\bar{X}_{1k} = \bar{X}_{2k} = X_k, k = 1, 2, \ldots, N$, $B_2 = B_1$, and $\bar{f} = f$.
2) Step 1: While $B_1 > B$ perform the following:
   - Set $k = 0$.
   - While $k \leq N$, perform the following:
     i) Set $c = 0$.
     ii) While $c \leq M$, perform the following:
        a) If $\sum_{i \in S_c^C \setminus k} \bar{X}_{1i} > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{i \in S_c^C \setminus k} Q_k \bar{X}_{1i}\}$.
        b) $c = c + 1$.
     iii) If $\bar{X}_{1k} = 1$, recompute its profit $\triangle f_k$ and cost b_k using the redefined d_c^C and E_c.
     iv) $k = k + 1$.
   - Remove the line extension j with smallest profit. That is, set $\bar{X}_{1j} = 0$.
   - $B_1 = B_1 - b_j$.
3) Step 2: Compute the profit $\bar{f}_1$ for the new solution $\bar{X}_{1k}, k = 1, 2, \ldots, N$. If $\bar{f}_1 > \bar{f}$, set $\bar{f} = \bar{f}_1$ and $\bar{X}_k = \bar{X}_{1k}, k = 1, 2, \ldots, N$.
4) Step 3: While $B_2 > B$ perform the following:
   - Set $k = 0$.
   - While $k \leq N$, perform the following:
     i) Set $c = 0$.
     ii) While $c \leq M$, perform the following:
        a) If $\sum_{i \in S_c^C \setminus k} \bar{X}_{1i} > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{i \in S_c^C \setminus k} Q_k \bar{X}_{1i}\}$.
        b) $c = c + 1$.
     iii) If $\bar{X}_{1k} = 1$, recompute its profit $\triangle f_k$ and cost b_k using the redefined d_c^C and E_c.
     iv) $k = k + 1$.
   - Remove the line extension j with smallest ratio of profit to cost. That is, set $\bar{X}_{2j} = 0$.
   - $B_2 = B_2 - b_j$.
5) Step 4: Compute the profit $\bar{f}_2$ for the new solution $\bar{X}_{2k}, k = 1, 2, \ldots, N$. If $\bar{f}_2 > \bar{f}$, then set $\bar{f} = \bar{f}_2$, and $\bar{X}_k = \bar{X}_{2k}, k = 1, 2, \ldots, N$.

APPENDIX B

LINE EXTENSIONS UNDER A CARDINALITY CONSTRAINT: PROBLEM P$_c$

The proof of the following hardness result uses a reduction from the Balanced Biclique Problem [36]. Since the proof uses standard techniques, we avoid stating it here explicitly for brevity.

Theorem 7: The decision problem corresponding to Problem P$_c$ is strongly NP-complete.The result above motivates the need for an efficient heuristic for Problem P$_c$ that can provide near-optimal solutions. This is our goal in the next section.

A. *Effective Heuristic for Problem P$_c$*

The idea behind the heuristic is similar to that for Problem P$_b$ (see Section IV-A). In Step 1 of the heuristic, we generate $L + 1$ problems as follows: Problem P^i, $i = 0, 1, \ldots, L$, uses the linear labor cost $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$. If the number of line extensions in the solution of Problem P^i is less than U, then Procedure IMPROVE-P$_c$ adds more line extensions to improve the profit; otherwise, if the number of line extensions in Problem P^i is greater than the upper bound U, Procedure REDUCE-P$_c$ removes some line extensions to satisfy the cardinality constraint. Let A_1 denote the best among these $L + 1$ feasible solutions. When the upper bound U is extremely small, this approach may not perform well because the number of extensions in the solution of Problem P^i without a cardinality constraint may far exceed the bound U. To address this special case, we include an additional step (Step 2): Choose one line extension k, $k = 1, 2, \ldots, N$, as an initial solution and then use Procedure IMPROVE-P$_c$ to enhance it. We, therefore, obtain N improved solutions. Let A_2 denote the best among these N improved solutions. The better of A_1 and A_2 is chosen as the final solution of the heuristic.

The heuristic is described below. The idea behind procedures IMPROVE-P$_c$ and REDUCE-P$_c$ is similar to that explained earlier for IMPROVE-RS (see Section III-C).

Heuristic H$_c$

Input: An instance of Problem P$_c$ with N products and M components. A positive integer L.

Output: A solution $Z_k, k = 1, 2, \ldots, N$, of Problem P$_c$ and the corresponding profit f.

1) Step 0 (initialization): $Z_k = 0, k = 1, 2, \ldots, N$; $f = 0$, and $i = 0$.
2) Step 1: While $i \leq L$, perform the following:
   - Set $l_{ci} = \frac{L-i}{L} l_c^h + \frac{i}{L} l_c^l$. Generate an instance of Problem P$_R$ by using the linear labor cost l_{ci}. Let the solution of this instance be $X_k^i, k = 1, 2, \ldots, N$.
   - Let $N^i = \sum_k X_k^i$. If $N^i < U$, use Procedure IMPROVE-P$_c$ to enhance the solution. Otherwise, if $N^i > U$, use Procedure REDUCE-P$_c$ to obtain a feasible solution. Let the improved solution be $\bar{X}_k^i, k = 1, 2, \ldots, N$, and its profit be $\bar{f}_i$. If $\bar{f}_i > f$, set $f = \bar{f}_i$ and $Z_k = \bar{X}_k^i, k = 1, 2, \ldots, N$.
   - Set $i = i + 1$.
3) Step 2: Set $i = 0$; While $i \leq N$, perform the following:
   - Set $X_i = 1$ and $X_j = 0, j \neq i$.
   - Use Procedure IMPROVE-P$_c$ to obtain a better solution. Let the improved solution be $\bar{X}_k^i, k = 1, 2, \ldots, N$, and its profit be $\bar{f}_i$. If $\bar{f}_i > f$, then set $f = \bar{f}_i$ and $Z_k = \bar{X}_k^i, k = 1, 2, \ldots, N$.
   - $i = i + 1$;
4) Step 3: Output the solution $Z_k, k = 1, 2, \ldots, N$, and its profit f.

Procedure IMPROVE-P$_c$

Input: A solution $X_k, k = 1, 2, \ldots, N$, of Problem P$_c$.

TABLE V
NOTATION FOR INSTANCES I, I_R^h, AND I_R^l, IN THE PROOF OF THEOREM 2

Instances	Component Labor Cost	Optimal Line Extension Set	Optimal Component Set	Optimal Profit	Revenue of the Optimal Line Extension Set	Cost of the Optimal Line Extension Set
I	l_c^h, l_c^l	S^*	$\mathcal{C}(S^*)$	f^*	–	–
I_R^h	l_c^h	S_h	$\mathcal{C}(S_h)$	f_h^h	–	–
I_R^l	l_c^l	S_l	$\mathcal{C}(S_l)$	f_l^l	φ_l	C_l

TABLE VI
NOTATION FOR THE DIFFERENT PROFITS IN THE PROOF OF THEOREM 2

Set of Line Extensions	Profit in Instance I	Profit in Instance I_R^h	Profit in Instance I_R^l
S_h	f_h	f_h^h	f_h^l
S_l	f_l	f_l^h	f_l^l

Output: An improved solution $\bar{X}_k, k = 1, 2, \ldots, N$, and its profit $\bar{f}$.

1) Step 0: Set $N_0 = \sum_k X_k$ and $\bar{X}_k = X_k, k = 1, 2, \ldots, N$.
2) Step 1: While $N_0 < U$, perform the following:
   - Set $c = 0$ and $k = 0$.
   - While $c \leq M$, perform the following:
     i) If $\sum_{k \in S_c^C} \bar{X}_k > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{k \in S_c^C} Q_k \bar{X}_k\}$.
     ii) $c = c + 1$.
   - While $k \leq N$, perform the following:
     i) If $\bar{X}_k = 0$, recompute its profit $\triangle f_k$ using the redefined values of d_c^C and E_c.
     ii) $k = k + 1$.
   - If there are unselected line extensions with positive profits, select the line extension j with largest profit, set $\bar{X}_j = 1$; otherwise, break.
   - $N_0 = N_0 + 1$.
3) Step 2: Compute the profit $\bar{f}$ corresponding to the solution $\bar{X}_k, k = 1, 2, \ldots, N$.

Procedure REDUCE-P_c

Input: An (infeasible) solution $X_k, k = 1, 2, \ldots, N$, of Problem P_c.

Output: A (feasible) solution $\bar{X}_k, k = 1, 2, \ldots, N$, and its profit $\bar{f}$.

1) Step 0: Set $N_0 = \sum_k X_k$, and $\bar{X}_k = X_k, k = 1, 2, \ldots, N$.
2) Step 1: While $N_0 > U$ perform the following:
   - Set $k = 0$.
   - While $k \leq N$, perform the following:
     i) Set $c = 0$. While $c \leq M$, perform the following:
        a) If $\sum_{i \in S_c^C \setminus k} \bar{X}_i > 0$, set $d_c^C = 0$ and $E_c = \max\{0, E_c - \sum_{i \in S_c^C \setminus k} Q_k \bar{X}_i\}$.
        b) $c = c + 1$.
     ii) If $\bar{X}_k = 1$, recompute its profit $\triangle f_k$ using the new d_c^C and E_c.
   - Remove line extension j with the smallest profit; i.e., set $\bar{X}_j = 0$.
   - $N_0 = N_0 - 1$.
3) Step 2: Compute the profit $\bar{f}$ corresponding to the solution $\bar{X}_k, k = 1, 2, \ldots, N$.

APPENDIX C
PROOFS OF THE TECHNICAL RESULTS

A) Proof of Theorem 1: The entries of the constraint matrix (6) are either 0, 1, or -1, with exactly one 1 and one -1 in each row. Thus, the constraint matrix is totally unimodular (see, e.g., [37]). Consequently, an optimal integral solution of Problem P_R can be obtained by solving its linear programming relaxation. The result follows.

B: Proof of Theorem 2: We first define additional notation for Instance I and its two variants (I_R^h and I_R^l) introduced in Section III-B.

1) Let S^* and $\mathcal{C}(S^*)$ be the optimal set of line extensions and the corresponding set of components for Instance I.
2) Recall that S_h is the optimal set of extensions for Instance I_R^h. Let $\mathcal{C}(S_h)$ be the corresponding set of components and f_h^h be the corresponding profit. For set S_h, let f_h^l be the profit computed using the lower labor cost l_c^l (as in Instance I_R^l).
3) Let S_l and $\mathcal{C}(S_l)$ be the optimal line extension set and component set and f_l^l be the optimal profit for Instance I_R^l. For set S_l, let f_l be the profit corresponding to the two-piece labor cost in the original Instance I and f_l^h be the profit computed using the higher labor cost l_c^h (as in Instance I_R^h).

Tables V and VI summarize the notation used in this proof. For set S_l, by the definition of μ and C_l, we have

$$C_l = \sum_{k \in S_l} \{o_k Q_k + d_k + g_k + \sum_{c \in S_k} m_c Q_k + \sum_{c \in S_k} l_c^l Q_k\} + \sum_{c \in \mathcal{C}(S_l)} d_c^C \geq \frac{1}{\mu} \sum_{k \in S_l} \sum_{c \in S_k} l_c^l Q_k. \quad (7)$$

Using $\varphi_l = (1+\beta)C_l$, we have

$$f_l^l = \varphi_l - C_l = \beta C_l. \tag{8}$$

Again, for set S_l, the profit f_l^h computed by using only the higher labor cost l_c^h is

$$f_l^h = \varphi_l - \Big\{ \sum_{k\in S_l} \Big\{ o_k Q_k + d_k + g_k + \sum_{c\in S_k} m_c Q_k + \sum_{c\in S_k} l_c^h Q_k \Big\} + \sum_{c\in \mathcal{C}(S_l)} d_c^C \Big\}.$$

Also, $l_c^l \geq (1-\alpha) l_c^h$. Thus, we have

$$\begin{aligned} f_l^h &\geq \varphi_l - \Big\{ \sum_{k\in S_l} \Big\{ o_k Q_k + d_k + g_k + \sum_{c\in S_k} m_c Q_k \Big\} + \sum_{c\in \mathcal{C}(S_l)} d_c^C \Big\} - \frac{1}{1-\alpha} \sum_{k\in S_l} \sum_{c\in S_k} l_c^l Q_k \\ &= \varphi_l - C_l - \frac{\alpha}{1-\alpha} \sum_{k\in S_l} \sum_{c\in S_k} l_c^l Q_k \\ &= f_l^l - \frac{\alpha}{1-\alpha} \sum_{k\in S_l} \sum_{c\in S_k} l_c^l Q_k. \end{aligned} \tag{9}$$

Using (7), (8), and (9), we have

$$\begin{aligned} f_l^l &\leq f_l^h + \frac{\alpha}{1-\alpha} \sum_{k\in S_l} \sum_{c\in S_k} l_c^l Q_k \leq f_l^h + \frac{\alpha\mu}{1-\alpha} C_l \\ &= f_l^h + \frac{\alpha\mu}{(1-\alpha)\beta} f_l^l. \end{aligned}$$

Therefore, when $\frac{\alpha\mu}{1-\alpha} < \beta$, we have

$$f_l^l \leq \frac{1}{1 - \dfrac{\alpha\mu}{(1-\alpha)\beta}} f_l^h. \tag{10}$$

For set S_h, f_h^h is the profit computed by using only the higher labor cost l_c^h, while f_h is its profit computed by using the two-piece labor cost l_c^h and l_c^l. Thus, we have

$$f_h^h \leq f_h. \tag{11}$$

For Instance I_R^h, f_l^h is the profit corresponding to set S_l, while f_h^h is the optimal profit. Therefore,

$$f_l^h \leq f_h^h. \tag{12}$$

For Instance I, f_h is the profit corresponding to set S_h, while f^* is the optimal profit. Then,

$$f_h \leq f^*. \tag{13}$$

Since f_l^l is the optimal profit for Instance I_R^l, where we use the lower labor cost l_c^l, while f^* is the optimal profit for Instance I, where we use the two-piece labor cost, we have

$$f^* \leq f_l^l. \tag{14}$$

Using (10)–(14), we get

$$f_l^h \leq f_h^h \leq f_h \leq f^* \leq f_l^l \leq \frac{1}{1 - \dfrac{\alpha\mu}{(1-\alpha)\beta}} f_l^h.$$

Thus, if $\dfrac{\alpha\mu}{1-\alpha} < \beta$, we have

$$\frac{f_h}{f^*} \geq \frac{f_l^h}{\frac{1}{1-\dfrac{\alpha\mu}{(1-\alpha)\beta}} f_l^h} = 1 - \frac{\alpha\mu}{(1-\alpha)\beta}.$$

The result follows.

C) Proof of Theorem 3: We first prove $S^* \subseteq S_l$ by contradiction. The relationship $S_h \subseteq S^*$ can be proved in a similar manner. Suppose $S^* \nsubseteq S_l$ and let $\tilde{S} = \{S^* \cap S_l\}$. Let $S = S^* \backslash \tilde{S} \neq \emptyset$. Consider the set of line extensions $\tilde{S}$ and let $\triangle f$ (respectively, $\triangle f_l$) be the incremental profit of adding set S to the set $\tilde{S}$ in Instance I (respectively, I_R^l). Since $S^* = S \cup \tilde{S}$ is an optimal set for Instance I, we have $\triangle f \geq 0$. Also, by definition, Instance I_R^l is identical to Instance I, except that the labor cost for each component c is the lower cost l_c^l instead of the original two-piece cost. Thus, we have $\triangle f_l > \triangle f \geq 0$. Now, let $\triangle f_l'$ be the incremental profit of adding set S to the set S_l in Instance I_R^l. Let $\mathcal{C}(\tilde{S})$ (respectively, $\mathcal{C}(S_l)$) be the set of components used by the extensions in $\tilde{S}$ (respectively, S_l). Since $\tilde{S} \subseteq S_l$, we have $\mathcal{C}(\tilde{S}) \subseteq \mathcal{C}(S_l)$. Therefore, it follows immediately that $\triangle f_l' \geq \triangle f_l \geq 0$. Thus, for Instance I_R^l, the profit offered by set S_l can be strictly improved by adding the line extensions in set S. This contradicts the assumed optimality of S_l for I_R^l. We conclude that $S^* \subseteq S_l$.

D) Proof of Theorem 6: We use the Partition Problem [36] for our reduction.

PARTITION

INSTANCE: A positive integer J and a set of positive integers $A = \{a_1, a_2, \ldots, a_n\}$ such that $\sum_{i=1}^n a_i = 2J$.

SOLUTION: Find a subset $A' \subseteq A$ such that $\sum_{a_i \in A'} a_i = \sum_{a_i \in A\backslash A'} a_i = J$.

Given an arbitrary instance of PARTITION, we construct the following instance of Problem P_b: each $a_i, i = 1, 2, \ldots, n$, corresponds to a distinct line extension, say k_i. Thus, the set of line extensions is $L = \{k_1, k_2, \ldots, k_n\}$. Each line extension uses the same component, say c. The other parameters are set as follows: $B = 2J$, $N = n$, $M = 1$, $\triangle R_k = 3a_k$, $Q_k = a_k$, $o_k = 1$, $d_k = 0$, $g_k = 0$, $d_c^C = J$, $m_c = 0$, $l_c^h = 0$, $l_c^l = 0$, and $E_c = 0$. For the constructed instance of Problem P_b, consider the following question:

DECISION PROBLEM: Does there exist a subset of extensions $L' \subseteq L$ with total profit $\phi \geq J$?

The decision problem is clearly in class NP. Also, our construction of the decision problem is in polynomial time in the size of the instance of PARTITION . We now show that the decision problem has an affirmative answer if and only if the PARTITION instance is satisfiable.

$\Longrightarrow$ Suppose the instance of PARTITION is satisfiable. Let $A' \subseteq A$ satisfy $\sum_{a_i \in A'} = J$. Then, introducing the line extensions

$L' = \{k_i : a_i \in A'\}$ provides the required profit J

$$\begin{aligned}\Phi &= \sum_{k \in L'} \{\Delta R_k - o_k Q_k - d_k - g_k\} - d_c^C \\ &\quad - \{(m_c + l_c^h) V_c^h + (m_c + l_c^l) V_c^l\} \\ &= 2 \sum_{a_i \in A'} a_i - J \\ &= J. \end{aligned} \tag{15}$$

$\Longleftarrow$ Suppose there exists $L' \subseteq L$ with the corresponding profit $\Phi \geq J$. Let $A' = \{a_i : k_i \in L'\}$. We show that $\sum_{a_i \in A'} a_i = J$. If $\sum_{a_i \in A'} a_i < J$, then by (15), the profit $\Phi = 2\sum_{a_i \in A'} a_i - J < J$, which contradicts the assumption $\Phi \geq J$. On the other hand, if $\sum_{a_i \in A'} a_i > J$, then the total cost is $\sum_{k \in L'} o_k Q_k + d_c^C = \sum_{a_i \in A'} a_i + J > 2J$, which contradicts the upper bound $2J$ on the available budget. Thus, $\sum_{a_i \in A'} a_i = J$. We, therefore, have the required partition of the set A.

References

[1] N. Giddens and A. Hofmann. (2010). Building your brand with brand line extensions. [Online]. Available: http://www.extension.iastate.edu/agdm/wholefarm/html/c5-52.html

[2] M. Marks,. (2008). How to develop and launch a new product. *HVACR Distribution Bus*. [Online]. Available: http://hvacrdistributionbusiness.com/mag/develop_ launch_ new

[3] H. S. Heese and J. M. Swaminathan, "Product line design with component commonality and cost-reduction effort," *Manuf. Serv. Oper. Manag.*, vol. 8, no. 2, pp. 206–219, 2006.

[4] W. J. Hopp and X. Xu, "Product line selection and pricing with modularity in design," *Manuf. Serv. Oper. Manag.*, vol. 7, no. 3, pp. 172–187, 2005.

[5] K. Ramdas, "Managing product variety: An integrative review and research directions," *Prod. Oper. Manag.*, vol. 12, no. 1, pp. 79–101, 2003.

[6] S. F. Brown. (2004). Toyota's global body shop. [Online]. Available: http://money.cnn.com/magazines/fortune/fortune_archive/2004/02/09/360102/index.htm

[7] G. Gardner. (1997). Production puzzle: Toyota rolls cars and minivans down the same line. [Online]. Available: http://business.highbeam.com/3355/article-1G1-19921572/production-puzzle-toyota-rolls-cars-and-minivans-down

[8] J. Parham. (2012). Line extension overload. *Progress. Grocer*. [Online]. Available: http://www.progressivegrocer.com/inprint/article/id5342/line-extension-overload

[9] J. Quelch and D. Kenny, Response titled "Extensions must be carefully planned and monitored" in the Perspectives on *The Logic of Product-Line Extensions, Harvard Business Review*, 1994.

[10] F. Geyer, "10 quick ways to drive organic growth," *Prophet—The Inspiratory*. [Online]. Available: http://www.prophet.com/theinspiratory/2013/12/05/10-quick-ways-to-drive-organic-growth

[11] J. Quelch and D. Harding, "Brands versus private labels: Fighting to win," *Harvard Business Review on Brand Management*, 1999, pp. 23–42.

[12] J. L. Mariotti, *The Complexity Crisis*. Avon, MA, USA: Adams Media, 2008.

[13] S. Balachander and S. Ghose, "Reciprocal spillover effects: A strategic benefit of brand extensions," *J. Market.*, vol. 67, no. 1, pp. 4–13, 2003.

[14] E. J. Nijssen, "Success factors of line extensions of fast-moving consumer goods," *Eur. J. Market.*, vol. 33, no. 5, pp. 450–474, 1999.

[15] S. K. Reddy, S. L. Holak, and S. Bhat, "To extend or not to extend: Success determinants of line extensions," *J. Market. Res.*, vol. 31, no. 2, pp. 243–262, 1994.

[16] C. H. Mason and G. R. Milne, "An approach for identifying cannibalization within product line extensions and multi-brand strategies," *J. Bus. Res.*, vol. 31, no. 2, pp. 163–170, 1994.

[17] D. Reiffen and M. R. Ward, "'Branded generics' as a strategy to limit cannibalization of pharmaceutical markets," *Manag. Decision Econ.*, vol. 28, pp. 251–265, 2007.

[18] S. Netessine and T. A. Taylor, "Product line design and production technology," *Market. Sci.*, vol. 26, no. 1, pp. 101–117, 2007.

[19] G. E. Fruchter, A. Fligler, and R. S. Winer, "Optimal product line design: Genetic algorithm approach to mitigate cannibalization," *J. Optim. Theory Appl.*, vol. 131, no. 2, pp. 227–244, 2006.

[20] K. Ramdas and M. S. Sawhney, "A cross-functional approach to evaluating multiple line extensions for assembled products," *Manag. Sci.*, vol. 47, no. 1, pp. 22–36, 2001.

[21] S. K. Fixson, "Modularity and commonality research: Past developments and future opportunities," *Concurrent Eng.: Res. Appl.*, vol. 15, no. 2, pp. 85–111, 2007.

[22] D. Robertson and K. Ulrich, "Planning for product platforms," *Sloan Manag. Rev.*, vol. 39, pp. 19–31, Summer 1998.

[23] F. Salvador, "Toward a product system modularity construct: Literature review and reconceptualization," *IEEE Trans. Eng. Manag.*, vol. 54, no. 2, pp. 219–240, May. 2007.

[24] S. Gupta and V. Krishnan, "Integrated component and supplier selection for a product family," *Prod. Oper. Manag.*, vol. 8, no. 2, pp. 163–181, 1999.

[25] S. W. Sanderson and M. Z. Uzumeri, *Managing Product Families.* Chicago, IL, USA: Irwin Press, 1996.

[26] M. H. Meyer and A. P. Lehnerd, *The Power of Product Platforms*. New York, NY, USA: Free Press, 1997.

[27] M. H. Meyer and J. H. Utterback, "The product family and the dynamics of core capability," *Sloan Manag. Rev.*, vol. 34, pp. 29–48, 1993.

[28] J. H. Mikkola and O. Gassmann, "Managing modularity of product architectures: Toward an integrated theory," *IEEE Trans. Eng. Manag.*, vol. 50, no. 2, pp. 204–218, May. 2003.

[29] T. Blecker and N. Abdelkafi, "The development of a component commonality metric for mass customization," *IEEE Trans. Eng. Manag.*, vol. 54, no. 1, pp. 70–85, Feb.. 2007.

[30] T. Agrawal, A. Sao, K. J. Fernandes, M. K. Tiwari, and D. Y. Kim, "A hybrid model of component sharing and platform modularity for optimal product family design," *Int. J. Prod. Res.*, vol. 51, no. 2, pp. 614–625, 2013.

[31] E. Labro, "The cost effects of component commonality: A literature review through a management-accounting lens," *Manuf. Serv. Oper. Manag.*, vol. 6, no. 4, pp. 358–367, 2004.

[32] J. Jiao, T. W. Simpson, and Z. Siddique, "Product family design and platform-based product development: A state-of-the-art review," *J. Intell. Manuf.*, vol. 18, pp. 5–29, 2007.

[33] R. Jans, Z. Degraeve, and L. Schepens, "Analysis of an industrial component commonality problem," *Eur. J. Oper. Res.*, vol. 186, pp. 801–811, 2008.

[34] R. Sanchez, and D. Sudharshan, "Learning by doing in the development of new products," *Market. Intell. Plann.*, vol. 11, no. 7, pp. 29–38, 1993.

[35] J. Hauser, G. J. Tellis, and A. Griffin, "Research on innovation: A review and agenda for marketing science," *Market. Sci.*, vol. 25, no. 6, pp. 687–717, 2006.

[36] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco, CA, USA: Freeman, 1979.

[37] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization.* Hoboken, NJ, USA: Wiley, 1988.

Liying Mu is currently working toward the Ph.D. degree in operations management at the Naveen Jindal School of Management, University of Texas at Dallas, Richardson, TX, USA.

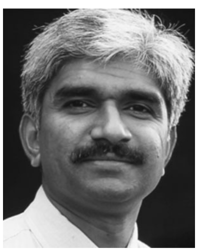
Her research interests include socially responsible operations management.

Dr. Mu is a Member of the Institute for Operations Research and the Management Sciences, and the Production and Operations Management Society.

Milind Dawande received the M.Sc. degree in computer science from the Birla Institute of Technology and Science, Pilani, India, and the M.S. and Ph.D. degrees in algorithms, combinatorics, and optimization from Carnegie Mellon University, Pittsburgh, PA, USA.

He is currently the Ashbel Smith Professor of Operations Management and the Area Coordinator of the operations management area at the Naveen Jindal School of Management, University of Texas at Dallas, Richardson, TX, USA.

His research interests include optimization theory and its applications in manufacturing, and service operations management.



Srinagesh Gavirneni received the B.Sc. degree in mechanical engineering from the Indian Institute of Technology, Chennai, India; the M.Sc. degree in industrial and manufacturing systems engineering from Iowa State University, Ames, IA, USA; and the Ph.D. degree in manufacturing and operations from Carnegie Mellon University, Pittsburgh, PA, USA.

He is currently an Associate Professor of Operations Management at the Johnson Graduate School of Management, Cornell University, Ithaca, NY, USA. During 2011–2012, he was a Visiting Associate Professor at the Nanyang Business School, Singapore. Previously, he was also an Assistant Professor at the Kelley School of Business, Indiana University, Bloomington, IN, USA; the Chief Algorithm Design Engineer at SmartOps, Pittsburgh; a Software Architect at Maxager Technology Inc., San Rafael, CA, USA; and a Research Scientist with Schlumberger, Houston, TX, USA. His research interests include supply chain management, inventory control, production scheduling, simulation, and optimization. His papers have appeared in *Management Science*, *Manufacturing & Service Operations Management*, *Operations Research*, *Production and Operations Management*, the *European Journal of Operational Research*, *Operations Research Letters*, *IIE Transactions*, and *Interfaces*.

Chelliah Sriskandarajah received the B.Sc. degree in mechanical engineering from the University of Sri Lanka, Katubedda, Sri Lanka; the M.S. degree in industrial engineering and management from the Asian Institute of Technology, Bangkok, Thailand; and the Ph.D. degree in production management from the National Polytechnic Institute of Grenoble, Grenoble, France.

He is currently a Professor of Operations Management and holds the Hugh Roy Cullen Chair at the Mays Business School, Texas A&M University, College Station, TX, USA. His research interests include the general area of supply chain management, logistics, production planning and scheduling, and performance evaluation of production/service systems. He has published more than 100 scholarly articles in leading international journals.

Mr. Sriskandarajah currently serves as a Departmental Editor of the *Production and Operations Management Journal*, and is a member of the Institute for Operations Research and the Management Sciences, and the Production and Operations Management Society (POMS). He served as the Associate Executive Director of POMS from March 2004 to April 2012. He was inducted as a POMS Fellow in April 2012.