Title	Robust Optimization Approach to Process Flexibility Designs with Contribution Margin Differentials
Authors	Wang, Shixin; Wang, Xuan; Zhang, Jiawei
Source	Manufacturing & Service Operations Management, 26 Oct 2020
Version	Accepted Version
DOI	10.1287/msom.2020.0913
Publisher	INFORMS
Copyright	© 2020 INFORMS

Robust Optimization Approach to Process Flexibility Designs with Contribution Margin Differentials

Shixin Wang
Department of Information, Operations & Management Sciences, Stern School of Business,
New York University, sw3333@nyu.edu

Xuan Wang
Department of Information Systems, Business Statistics and Operations Management,
Hong Kong University of Science and Technology, xuanwang@ust.hk

Jiawei Zhang
Department of Information, Operations & Management Sciences, Stern School of Business,
New York University, jzhang@stern.nyu.edu

Problem definition: The theoretical investigation of the effectiveness of limited flexibility has mainly focused on the performance metric that is based on the maximum sales in units. However, this could lead to substantial profit losses when the maximum sales metric is used to guide flexibility designs whereas the products have considerably large profit margin differences. *Academic/practical relevance:* We address this issue by introducing margin differentials into the analysis of process flexibility designs, and our results can provide useful guidelines for the evaluation and design of flexibility configurations when the products have heterogeneous margins. *Methodology:* We adopt a robust optimization framework and study process flexibility designs from the worst-case perspective by introducing the Dual Margin Group Index (DMGI). *Results and Managerial Implications:* We show that a general class of worst-case performance measures can be expressed as functions of a design's DMGIs and the given uncertainty set. Moreover, the DMGIs lead to a partial ordering that enables us to compare the worst-case performance of different designs. Applying these results, we prove that under the so-called part-wise independently symmetric uncertainty sets and a broad class of worst-case performance measures, the alternate long chain design is optimal among all long chain designs with equal number of high profit products and low profit products. Finally, we develop a heuristic based on the DMGIs to generate effective flexibility designs when products exhibit margin differentials.

Key words: Process flexibility; Flexible Production; Profit maximization; Robust Optimization.

---

1. Introduction

In today's competitive business environment where consumers' demand have become increasingly uncertain and volatile, the ability to match supply and demand is arguably key to a firm's success in many industries. Unfortunately, the capacity investment decisions are typically made far in advance before the uncertain demand is realized, which usually leads to inevitable misalignments between supply and demand with potential severe economic consequences. In a well-known example (Greenberg 2001), in year 2000, Chrysler faced higher than expected demand for their newly launched PT Cruiser, but experienced dampened demand for their hitherto well-selling Town and Country minivan. However, since both products were manufactured on their respective dedicated production lines, Chrysler could not use the plant with spare capacity to produce the PT Cruiser, and this misalignment between supply and demand cost the company approximately $2 billion.

Intuitively, this supply-demand mismatch could have been significantly reduced if Chrysler had adopted a *flexible* production system that is able to produce both models, which would enable the firm to shift production from one plant to another plant in accordance with the realized demand. Fortunately, most of the automobile giants have recognized the competitive advantage of equipping their manufacturing plants with a flexible production system (e.g., Phelan 2002, Boudette 2006). For example, General Motors, which has built in more flexibility across its production facilities, was able to successfully meet soaring customer demand for the Equinox and GMC Terrain by increasing the production volume of these two models by more than 60,000 to 80,000 more vehicles using the underutilized capacity of their Ingersoll and Oshawa plants in 2010 (Chevrolet 2010).

Process flexibility, defined as a firm's ability to "build different types of products in the same manufacturing plant or on the same production line at the same time" (Jordan and Graves 1995, p. 577), has been proven to be a successful operational strategy in manufacturing industries to hedge against demand uncertainty and volatility. A seminal work by Jordan and Graves (1995) demonstrates that even just a little flexibility, if configured in the right way, can be extremely effective in mitigating supply-demand mismatch. Among all the designs with limited flexibility, the *chaining strategy*, and in particular the *long chain design* proposed by Jordan and Graves (1995), is perhaps one of the most influential strategies both studied in the literature and used in practice. They show that there exist sparse networks with limited flexibility that perform almost as well as full flexibility, and this limited flexibility accrues the greatest benefits when configured to chain products and plants together to the greatest extent possible. Motivated by the findings in Jordan and Graves (1995), the effectiveness of the long chain and designs with limited flexibility has been investigated theoretically in many recent works, such as Chou et al. (2010b), Bassamboo et al. (2010), Chou et al. (2011), Simchi-Levi and Wei (2012, 2015), Wang and Zhang (2015), Chen et al. (2015) and Désir et al. (2016). We refer the readers to Wang et al. (2019) for a recent survey.

The existing literature on process flexibility designs has largely focused on performance metrics based on the maximum sales in units. More specifically, for any realized demand instance, one solves a max-flow problem to determine the optimal allocation of the available capacities to satisfy the demand as much as possible. However, many production systems in practice exhibit profit margin differentials which may come from the demand side and/or the supply side. For example, products may differ in their selling prices and securing a unit capacity at different plants may incur different costs. In case there exist margin differentials, the various design principles and heuristics developed in the existing literature under the maximum sales metric may not be effective any more and could potentially result in substantial profit loss. It has been shown in a seminal work by Van Mieghem (1998) that margin differentials are an essential element that affects the value of flexibility. The author considers a two-product firm with the choice of investing in product-dedicated resources and/or a fully flexible resource, and demonstrates that margin differentials can significantly affect on the value of flexibility. In particular, Van Mieghem (1998) shows that contrary to the intuitive belief that flexible capacity provides no additional value when product demands are perfectly positively correlated, it can be optimal to invest in flexible capacity even with perfectly positively correlated product demands when there exists a positive margin differential. The underlying logic here is that, besides its ability to appropriately adjust capacity allocation to the demand realizations, flexibility provides an additional opportunity to improve revenues by producing more of highly profitable products at the expense of less profitable products.

Our objective in this paper is to study effective flexibility designs when there exist margin differentials among the products, and the evaluation metric that aims to maximize the total profit is of interest. Different from the works by Fine and Freund (1990) and Van Mieghem (1998) which examine the optimal *capacity investment strategy* in dedicated resources and/or a *fully* flexible resource, our focus is on the *flexibility configuration design* where the plant capacities are exogenously given. Without assuming full distributional information of the random product demands, we adopt a robust optimization approach to examine the worst-case performance of flexibility designs. In the remainder of this section, we first summarize our main results and contributions, and present a brief literature review in Section 1.2.

1.1. Main Results and Contributions

In Section 3, we introduce the Dual Margin Group Index (DMGI) to study the worst-case performance of flexibility designs when there exist margin differentials among the products. The name of this index comes from the fact that it uses the dual variable information of an optimization problem which maximizes the total profit. Intuitively speaking, this dual variable represents the shadow price associated with a product, which measures the "marginal profitability level" of a product's demand. For a fully flexible network, it has the best capability to route plant capacities to the appropriate products with the best profit potential and achieve the maximum total profit. However, for a design with limited flexibility, a group of products and a subset of plants may form a "bottleneck" that blocks the network from the best usage of the available plant capacities to achieve the maximum profit potential. The DMGI associated with a sparse flexibility design captures such a bottleneck in the worst-case under a given set of dual price vectors. We show that the DMGIs can be used to characterize a broad class of worst-case performance measures for general unbalanced and asymmetric production systems, which in turn leads to a partial ordering[1] that enables us to compare the performance of different flexibility designs.

Applying the above results, we show in Section 4 that a class of long chain designs, the *alternate long chain*, is optimal among all long chain designs where there are two product categories and the

[1] A relation "$\leq$" is a *partial order* on a set S if it has: (1) Reflexivity: $a \leq a$ for all $a \in S$; (2) Antisymmetry: $a \leq b$ and $b \leq a$ implies $a = b$; (3) Transitivity: $a \leq b$ and $b \leq c$ implies $a \leq c$.

number of high margin products is equal to that of low margin products. Intuitively, among all the long chain designs with equal number of high profit and low profit products, the alternate long chain has the best potential to fulfill the demand from high profit products compared with designs where multiple high profit products share a common plant. One implication of the above optimality result is that, *within* the class of long chain production systems (which has been a well-known class of designs for its effective performance), it is most beneficial to evenly "spread out" the products from the two product groups in an alternate manner and to equip each plant to be able to produce both high profit and low profit products. As a result, the high profit products never have to directly compete for shared resources. Our comprehensive numerical results confirm that this intuition is also valid when the products have more than two profit margins in a connected design. Intuitively, when the demands of all the products cannot be fully satisfied due to capacity constraints and/or limited production flexibility, one would prefer to prioritize the demand fulfilment for higher margin products. To achieve this, it is more favorable to configure flexibility in such a way that enables each plant to produce both high margin and low margin products.

However, the *direct* competition between high margin products would become a less important issue if their demands are negatively correlated. In this case, it may be more beneficial to pool the negatively correlated products and serve them by the same resources. In fact, we show that the alternate long chain is not necessarily optimal among the class of *2-flexibility designs* with equal number of high profit and low profit products, which is in contrast with the result by Simchi-Levi and Wei (2015) that establishes the optimality of the long chain design among 2-flexibility designs under symmetric uncertainty sets in the absence of margin differentials. A particular example of a 2-flexibility design, the so-called *disjoint long chain*, is composed of *disconnected* shorter chains where products in one component all have high margin and the other component consists of products with low margin, can achieve a better worst-case performance than the alternate long chain when the demands of products *within* each group are *negatively correlated*. The disconnectivity of the disjoint long chain makes half of the total capacity exclusive for the high margin products, and the negative correlation among the high margin product demands can make the half total capacity that are exclusive to them (which are not shared by any low margin products either directly or indirectly) well utilized since the "competition" for shared resources among the high margin products would not be an issue due to the negative correlation. Therefore, our results suggest that when there exist margin differentials and in the case of negatively correlated demands within each product group, it may be more beneficial to construct disconnected shorter chains for each group to reduce competition across groups. Here we would like to point out that, although the alternate long chain is no longer optimal among the class of 2-flexibility designs, our numerical results suggest that the performance of designs guided by the above "spread out" insight compare extremely well with the performance of the optimal configuration even if the demands of products within each group are negatively correlated. This suggests the robustness of the "spread out" insight and may provide useful design guidelines for generating effective flexibility designs in the presence of profit margin differentials. In particular, when the demands of the higher margin products do not exhibit strong negative correlation and/or the correlation information is not known, it would be favorable to follow the "spread out" insight for effective flexibility designs. In case there exists strong negative correlation among products within the same group, one may consider partitioning the products into clusters such that products with the same profit margin share the same set of resources when constructing flexibility designs (we refer to Appendix B in Wang et al. (2020) for a more detailed discussion[2]), or one may resort to other approaches (such as SAA) to solve for the optimal design.

The above "spread out" insight stems from the optimality of the alternate long chain among all long chain designs with equal number of high profit and low profit products, which was proved by using Theorem 2 that one flexibility design has a better worst-case performance than another if the DMGI values of the former are larger than the latter. However, the level of "spread-out" of a flexibility design can sometimes be difficult to quantify. These motivate us to develop an implementable

[2] To meet the journal's page limit requirements, we have relegated some numerical studies and appendices in a peer-reviewed, unabridged version of the manuscript Wang et al. (2020).

heuristic for generating effective sparse flexibility designs in the presence of margin differentials based on a *quantifiable* index, where in each iteration the heuristic aims to help increase the value of DMGIs. Our computational study in Section 5 suggests that the sparse design generated by our DMGI-Heuristic can capture most of the benefit of full flexibility from the expected profit point of view, and achieves satisfactory performance in the worst case as well for small to moderate level of demand uncertainty. To better understand the effectiveness of our DMGI-Heuristic, we have also compared its performance with other algorithms proposed in the literature which do not consider margin differentials among the products. Our benchmarking algorithms include the Plant Cover Index (PCI) Heuristic by Simchi-Levi and Wei (2015), the Expander heuristic developed in Chou et al. (2011), the chaining structures, the Sample Average Approximation (SAA) method, the stochastic optimization based heuristic proposed by Feng et al. (2017) (denoted as MDEP-Heuristic), the Thresholded Probabilistic Construction (TPC) heuristic by Chen et al. (2019) and the Generalized Chaining Gap (GCG) based heuristic developed by Shi et al. (2019). Compared with the MDEP-Heuristic and SAA, our DMGI-Heuristic offers comparable performance when the actual demand follows the original distribution. In case that there exists demand distribution mis-specification where the demands across products are assumed to be independent but the actual demands are positively correlated within each product group, our DMGI-Heuristic offers similar performance in terms of the average performance and achieves better worst-case performance than the MDEP-Heuristic and the SAA approach. For the comparison with the PCI-Heuristic, the Expander heuristic, the chaining structures, the TPC heuristic and the GCG heuristic, we have conducted numerical experiments with identical profit margins, as these benchmark heuristics are developed with the objective of maximizing total sales in units. We compare our DMGI-Heuristic with the PCI-Heuristic, the Expander heuristic and the chaining structures under a balanced and symmetric system. The numerical results suggest that our DMGI-Heuristic is comparable to PCI-Heuristic in terms of average performance, and achieves a slightly better worst-case performance than that by PCI-Heuristic. For the Expander algorithm and the chaining structures, our DMGI-Heuristic consistently achieves a better performance from both the average and the worst-case perspective, and the improvement in the worst-case ratio becomes more significant as the demand becomes more volatile. We have further compared our DMGI-Heuristic with the TPC, GCG and Expander heuristics under unbalanced and asymmetric systems. Numerical results show that our DMGI-Heuristic offers comparable performance to the Expander heuristic. As for the GCG heuristic, our DMGI-Heuristic compares well with GCG when the actual demand follows the original distribution, and achieves considerable improvement over GCG in terms of worst-case performance when the actual demands are positively correlated within groups but assumed to be independent. For the TPC heuristic, our DMGI-Heuristic achieves a better performance consistently. Finally, we summarize and conclude the paper in Section 6 with directions for future research.

1.2. Literature Review

In this section, we briefly review literature that is most relevant to our work. Following the work by Jordan and Graves (1995), the effectiveness of the long chain and designs with limited flexibility has been justified theoretically in many recent works. Chou et al. (2010b) is among the first to provide theoretical justification of the effectiveness of the long chain and sparse structures. In particular, Chou et al. (2010b) develop a method to compute the expected demand satisfied by the long chain design in an asymptotically large system, and show that the expected sales achieved by the long chain is very close to that of full flexibility under some common demand distributions. Wang and Zhang (2015) adopt a distributionally robust approach to study the performance of flexibility designs and obtain an asymptotic, distribution-free lower bound on the ratio between the expected sales of the long chain relative to that of full flexibility. For finite size systems, Simchi-Levi and Wei (2012) develop a decomposition technique to analyze the expected sales achieved by the long chain and prove that the long chain design is optimal among 2-chains, i.e., all designs in which each product node and each plant node are incident to exactly two arcs. However, if we relax the class of 2-chains to all designs with $2n$ arcs, a recent result by Désir et al. (2016) shows that the long chain is no longer optimal among all designs with $2n$ arcs. Chou et al. (2011) prove that there exists a sparse graph which achieves $(1-\epsilon)$-optimality of the fully flexible structure in the worst-case demand scenario. More specifically, they show that when the demand and capacity levels are identical and balanced and demands are bounded by a constant λ, then there exists an $(\alpha, \lambda, \Delta)$-expander which performs within $(1-\alpha\lambda)$-optimality of the fully flexible system for *every* demand scenario. More recently, Chen et al. (2015) introduce probabilistic graph expanders and prove that in a balanced and symmetric system with n plants and n products, there exists a probabilistic expander with $O(n \ln(1/\epsilon))$ arcs that guarantees $(1-\epsilon)$-optimality with high probability. A follow-up work by Chen et al. (2019) generalizes the results in Chen et al. (2015) to unbalanced and asymmetric systems. Most of the above works consider offline allocation decisions, in which case how to allocate available capacities to fulfill demand is decided after demand realization. A recent work by Asadpour et al. (2020) investigates the effectiveness of the long chain design in the case where the allocation decisions must be made in an online fashion, and show that the long chain is still extremely effective in reducing supply-demand mismatch in the online decision-making setting. Simchi-Levi et al. (2018) incorporate inventory decisions with process flexibility for supply chain risk mitigation. They propose a robust optimization formulation where the first-stage decision variables are inventory levels and the second-stage decision variables are production quantities. They complement the existing literature on 2-chain by showing that increasing the degree of flexibility beyond 2-chain can still achieve significant benefit.

The most closely related to our work is the paper by Simchi-Levi and Wei (2015), which introduces an index that only depends on the design structure, the plant cover index, to study the worst-case performances of flexibility designs. They prove that if all of the plant cover indices of one flexibility design, $\mathcal{A}_1$, are greater than or equal to the plant cover indices of another design, $\mathcal{A}_2$, then $\mathcal{A}_1$ is more robust than $\mathcal{A}_2$ under any symmetric uncertainty set and a broad class of performance functions. Other indices have also been developed in the literature to compare the effectiveness of different flexibility designs, such as the original index developed in Jordan and Graves (1995), the structural flexibility index in Iravani et al. (2005), WS-APL index in Iravani et al. (2007), and the expansion index in Chou et al. (2008).

Research on flexibility configuration designs that explicitly takes margin differentials into account has received much less attention in the operations literature. Chou et al. (2010a) explicitly distinguish between primary and secondary links, where primary links connect plants with the products they are originally designed to manufacture and are associated with a lower cost than those of the secondary links. The authors obtain asymptotic lower bounds on the chaining efficiency, which captures the relative improvement a sparse design can achieve over a dedicated design compared with full flexibility. Their results show that chaining can be less beneficial when the efficiency loss associated with secondary arcs is taken into consideration. Mak and Shen (2009) adopted a stochastic optimization based approach to study the problem of flexibility design problems with price and cost differentials incorporated into the model. The authors use a Lagrangian-based approach with state-independent multipliers to directly search for an optimal flexibility configuration that maximizes the system's expected total profits. Their computational results suggest that their solution can be significantly better than the chaining strategy when the cost parameters are heterogeneous. More recently, Feng et al. (2017) extended the model of Mak and Shen (2009) by introducing an additional set of parameters on the unit capacity consumption rate of each plant-product pair to capture heterogeneous production efficiency, and the authors have developed an efficient solution algorithm based on sample average approximation to solve the extended model.

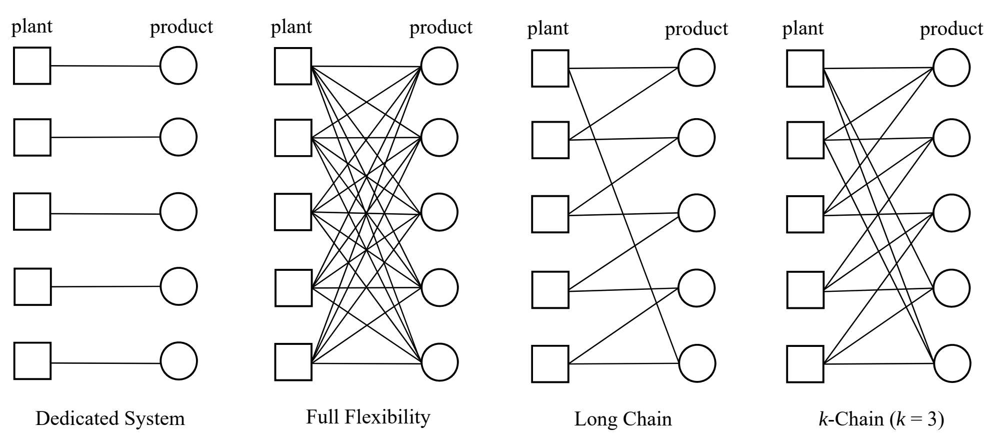
Process flexibility has also been studied in various application settings. Iravani et al. (2007) and Wallace and Whitt (2005) study process flexibility in a call-center limited labor cross-training setting. Graves and Tomlin (2003) analyze the benefits of flexibility in multi-stage supply chains. Bassamboo et al. (2010) study the optimal flexibility configurations in the application of resource portfolio investment. Tsitsiklis and Xu (2017) analyze the benefits of limited flexibility and resource pooling in multi-server queueing networks. We refer readers to the survey by Chou et al. (2008) and Buzacott and Mandelbaum (2008) for a more detailed review on process flexibility design.

2. Model

We consider a system with m plants and n products for some arbitrarily fixed positive integers m and n. Let $A := \{a_1, ..., a_m\}$ denote the set of plant nodes, where plant j has a fixed capacity C_j for each $1 \leq j \leq m$. Denote $B := \{b_1, ..., b_n\}$ as the set of product nodes, where for each $1 \leq i \leq n$, the profit margin of product i is equal to p_i. In our model, we assume that there is a total of T different margins $P_1 < P_2 < \cdots < P_T$, and $p_i \in \{P_1, P_2, ..., P_T\}$ for each $1 \leq i \leq n$. Let S_t denote the subset of products with profit margin P_t, i.e., $S_t = \{b_i \in B : p_i = P_t\}$. A system is said to be *balanced* if the numbers of plants and products are equal. A system is called *symmetric* if all plants have the same capacity and all products have independent and identically distributed (i.i.d.) demands.

A flexibility design $\mathcal{A}$ is represented by the arc set of a bipartite graph defined on sets A and B, where an arc from plant node a_j to product node b_i means that plant j is capable of producing product i, and we denote $|\mathcal{A}|$ as the number of edges in flexibility design $\mathcal{A}$. For example, the *full flexibility design* in which each plant has the capability to produce all the products is represented as $\mathcal{F} = \{(a_j, b_i) | 1 \leq i \leq m, 1 \leq j \leq n\}$. In a balanced system, we define the *long chain* design, where each product can be produced by exactly two adjacent plants, as $\mathcal{L} = \{(a_i, b_i), (a_i, b_{i+1}) | i = 1, ..., n-1\} \cup \{(a_n, b_n), (a_n, b_1)\}$; and the *dedicated design*, where each product can be produced by exactly one plant, as $\mathcal{D} = \{(a_i, b_i) | 1 \leq i \leq n\}$. See Figure 1 for an illustration of the long chain, together with dedicated and fully flexible systems when $n = 5$.

Figure 1 **Flexibility designs.**



Given a demand instance $\boldsymbol{d} = [d_1, ..., d_n]$, the *maximum profit* that can be achieved by a flexibility design $\mathcal{A}$, denoted by $g(\mathcal{A}, \boldsymbol{d})$, can be obtained by solving the following linear program (LP), where the decision variables x_{ij} represent the amount of demand for product i fulfilled by plant j:

$$\begin{aligned} g(\mathcal{A}, \boldsymbol{d}) = \max & \sum_{(a_j,b_i)\in\mathcal{A}} p_i x_{ij} \\ \text{s.t.} \quad & \sum_{j=1}^{m} x_{ij} \le d_i, \quad \forall 1 \le i \le n, \\ & \sum_{i=1}^{n} x_{ij} \le C_j, \quad \forall 1 \le j \le m \\ & x_{ij} \ge 0, \quad \forall 1 \le i \le n, 1 \le j \le m. \end{aligned} \tag{1}$$

It is appropriate to point out that in our model, the contribution margin is product dependent. It would be ideal to allow the contribution margins to be plant-product pair specific, but this would make the analysis much more challenging under our current solution approach and we shall leave it to future research. We would also like to remark that our model has implicitly assumed that the firm is a price taker since the profit margins are not affected by the production quantities chosen.

In the process flexibility literature, a very popular metric for evaluating the performance of flexibility designs has been based on the *maximum sales* in *units*. More specifically, one solves a max flow problem that is similar to (1) with the objective function replaced by $\max \sum_{(a_j,b_i)\in\mathcal{A}} x_{ij}$ to compute the maximum sales achieved by flexibility design $\mathcal{A}$ under a given demand instance $\boldsymbol{d}$. However, the maximum sales metric does not take into account any potential margin differentials that may appear among the products, which could lead to substantial profit losses when the maximum sales metric is used to guide flexibility designs whereas the products have considerably large margin differences. Our objective in this paper is to examine the performance of flexibility designs under the maximum profit metric. To the best of our knowledge, we are among the first to explicitly consider product margin differentials in flexibility designs.

2.1. Notation

Let $\mathbb{R}_+$ be the set of non-negative reals. Let $\mathbb{R}_+^n$ denote the n-dimensional vector space of non-negative reals. Let $\mathbb{Z}_+$ be the set of non-negative integers. Denote $\mathbb{1}\{\cdot\}$ as the indicator function. For a vector $\boldsymbol{x} \in \mathbb{R}^n$, denote $\min^i(\boldsymbol{x})$ as the i-th smallest element in the set $\{x_1, ..., x_n\}$. Denote $[n]$

as the set of integers from 1 to n, and let $\Omega([n])$ denote the set of all permutations of $[n]$. For any demand instance $\boldsymbol{d} \in \mathbb{R}_+^n$, define $\boldsymbol{d}_\sigma := [d_{\sigma(1)}, ..., d_{\sigma(n)}]^T$ for any index permutation $\sigma \in \Omega([n])$.

For flexibility design $\mathcal{A}$ with bipartition sets A and B, define $N(u, \mathcal{A}) = \{v \,|\, (u,v) \in \mathcal{A} \text{ or } (v,u) \in \mathcal{A}\}$ for any $u \in A \cup B$. That is, $N(u, \mathcal{A})$ is the set of neighboring nodes of u in the bipartite graph defined by $(A, B, \mathcal{A})$. Moreover, for any subset $S \subseteq A$ or $S \subseteq B$, let $N(S, \mathcal{A}) = \bigcup_{u \in S} N(u, \mathcal{A})$.

2.2. Robust (Worst-Case) Measures

For a given demand instance $\boldsymbol{d}$, a deterministic measure is a function that maps $\boldsymbol{d}$ and a flexibility design $\mathcal{A}$ to a real number. For example, the maximum profit achieved by a flexibility design, $g(\cdot)$, is one such deterministic measure. Given a deterministic measure function f, we denote $R^f(\cdot)$ as the robust measure of f, which is defined as the worst-case performance among all possible demand instances in a given uncertainty set U:

$$R^f(\mathcal{A}, U) := \min_{\boldsymbol{d} \in U} f(\boldsymbol{d}, \mathcal{A}).$$

In other words, $R^f(\cdot)$ is a robust counterpart of f that maps a flexibility design $\mathcal{A}$ and an uncertainty set U to a real number, which considers the worst-case performance and measures the "robustness" of $\mathcal{A}$ under U. Since the product demand can never be negative, we assume that any uncertainty set U considered in this paper lies in $\mathbb{R}_+^n$. We also assume that any deterministic measure function f is continuous in $\boldsymbol{d}$, which ensures that its robust counterpart $R^f(\cdot)$ is always well defined.

Let $s_1, ..., s_T$ be the index sets associated respectively with the T product categories such that $s_1 \cup \cdots \cup s_T = [n]$ and $i \in s_t$ if and only if $b_i \in S_t$, and denote $-s_t := [n] \setminus s_t$. For a given demand instance $\boldsymbol{d}$, let $\boldsymbol{d}_{s_t} := \{d_i \,|\, b_i \in S_t\}$ be the demand vector with appropriate dimension that is associated with products in subset S_t, where all the products in the same category S_t have identical margin P_t. Similarly, define $\boldsymbol{d}_{-s_t} := \{d_i \,|\, b_i \notin S_t\}$. For each product category S_t, we use $\sum \boldsymbol{d}_{s_t} := \sum_{i \in s_t} d_i$ to represent the total demand in product category S_t under demand instance $\boldsymbol{d}$. A deterministic measure function $f(\cdot)$ is said to be *monotonic* in profit under fixed total demand in each product category $\sum \boldsymbol{d}_{s_t}$ and separable in $\sum \boldsymbol{d}_{s_t}$ for each $1 \le t \le T$, if there exists some function h such that

$$f(\mathcal{A}, \boldsymbol{d}) = h\left(g(\mathcal{A}, \boldsymbol{d}), \sum \boldsymbol{d}_{s_1}, ..., \sum \boldsymbol{d}_{s_T}\right), \tag{2}$$

and $h(x, y_1, ..., y_T)$ is strictly increasing in x for any fixed real numbers $y_1, ..., y_T$. Let Γ denote the set of all robust measures associated with deterministic measure functions that are monotonic in profit under fixed total demand in each product category $\sum \boldsymbol{d}_{s_t}$ and separable in $\sum \boldsymbol{d}_{s_t}$ for each $1 \le t \le T$. It is easy to check that most of the commonly used deterministic measure functions satisfy the above condition and their robust counterparts belong to Γ. Examples include $f(\mathcal{A}, \boldsymbol{d}) = g(\mathcal{A}, \boldsymbol{d})$ (the maximum profit), $f(\mathcal{A}, \boldsymbol{d}) = g(\mathcal{A}, \boldsymbol{d}) - \sum_{i=1}^{n} p_i d_i$ (the potential profit loss due to capacity constraint), $f(\mathcal{A}, \boldsymbol{d}) = g(\mathcal{A}, \boldsymbol{d}) - g(\mathcal{F}, \boldsymbol{d})$ (the profit gap between full flexibility and $\mathcal{A}$).

2.3. Demand Uncertainty Sets

One of the major modelling decisions that we have deferred until this point is the choice of uncertainty sets to describe possible demand scenarios. In the worst-case analysis, symmetric uncertainty sets are frequently used to model symmetric demand variations, where a set U is said to be *symmetric* if for any $\boldsymbol{d} \in U$, $\boldsymbol{d}^{\sigma} \in U$ for any permutation $\sigma \in \Omega([n])$ (see Simchi-Levi and Wei 2015 for examples of symmetric uncertainty sets). Symmetric uncertainty sets imply that the worst-case performance will not change if we relabel the products. A generalization of symmetric uncertainty sets is the class of *symmetric perturbation uncertainty sets*, where a set U is called a symmetric perturbation uncertainty set if $E := \{\boldsymbol{x} - \boldsymbol{\mu} \,|\, \boldsymbol{x} \in U\}$ is symmetric for some fixed $\boldsymbol{\mu}$. Intuitively, one can interpret symmetric perturbation uncertainty sets as having product demands estimated to be $\boldsymbol{\mu}$, and the estimation error (perturbation) has the same fluctuation across products around $\boldsymbol{\mu}$.

In our model, products are grouped into T categories $S_1, ..., S_T$, where products in the same category have identical margin. In view of the (margin) heterogeneity across different product categories, we consider a class of uncertainty sets that do not require the estimation error around $\boldsymbol{\mu}$ to have the same fluctuation across *all* products, but instead only assume that the perturbation set is symmetric within each product category. Moreover, we assume independent demand across different product groups in the sense that the possible demand scenarios of products in S_t are independent of the values of $\boldsymbol{d}_{-s_t}$ and hence uncertainty set U can be written as the Cartesian product $U = U_1 \times \cdots \times U_T$, where $U_t := \{\boldsymbol{d}_{s_t} \,|\, \boldsymbol{d} \in U\}$. Formally, let $\Omega(s_t)$ denote the set of all permutations of index set s_t of the products in S_t. We have the following definitions:

Definition 2.1 (Part-wise Independently Symmetric Uncertainty Set) *An uncertainty set $U \subset \mathbb{R}^n_+$ is* part-wise independently symmetric *if (1) $U = U_1 \times \cdots \times U_T$, and (2) for any $\boldsymbol{d}_{s_t,-s_t} \in U$, we have $\boldsymbol{d}_{\sigma^t(s_t),-s_t} \in U$ for any permutation $\sigma^t \in \Omega(s_t)$ for each $1 \leq t \leq T$.*

Definition 2.2 (Part-wise Independently Symmetric Perturbation Uncertainty Set) *An uncertainty set $U \subset \mathbb{R}^n_+$ is a* part-wise independently symmetric perturbation uncertainty set *if $E := \{\boldsymbol{d} - \boldsymbol{\mu} \,|\, \boldsymbol{d} \in U\}$ is part-wise independently symmetric for some fixed $\boldsymbol{\mu}$.*

It is worth pointing out that the uncertainty sets considered in Definition 2.2 are more general than that in Definition 2.1, since U is independently symmetric if and only if U is independently symmetric around some $\boldsymbol{\mu}$ with $\mu_i = \mu_j$ for all $1 \leq i, j \leq n$. Here we would like to remark that the goal of the current paper is to develop tools that can identify flexibility designs that perform well for a general class of uncertainty sets, rather than constructing an uncertainty set to model demand uncertainty. To the best of our understanding, in the robust optimization literature, the uncertainty sets are chosen to balance the computational and/or analytical tractability of the resultant problem formulation with the quality of the solution. Different uncertainty sets may work well for different models, and we refer the readers to Bertsimas et al. (2011) for a list of most commonly used uncertainty sets. As we shall show later, the particular choice of the part-wise independently symmetric perturbation uncertainty sets allows us to derive an index that can identify flexibility designs that perform well.[3] Unfortunately however, we have not been able to derive a similar index for more general uncertainty sets.

3. Robust Measures and the Dual Margin Group Index

In this section, we formulate the robust measure to evaluate the worst-case performance of general unbalanced and asymmetric flexibility designs when there exist margin differentials among the products. We first define an index, the so-called Dual Margin Group Index (DMGI), which can be used to characterize the worst-case performance of a flexibility design as we show in later sections.

[3] In Wang et al. (2020) Appendix C, we provide several examples that illustrate the applicability of our chosen uncertainty sets.

3.1. Definition of the Dual Margin Group Index

In this subsection, we formally define our Dual Margin Group Index (DMGI) that will become useful for evaluating and comparing the worst-case performance of flexibility designs. Our DMGI is partly inspired by the Plant Cover Index (PCI) introduced by Simchi-Levi and Wei (2015), which considers a system where all the n products have identical profit. For any integer $k \in \{0, 1, ..., n\}$, the PCI at k is defined as the minimum plant capacity required to create a vertex cover on a flexibility design, given that the vertex cover contains exactly k product nodes. We next introduce our DMGI to account for margin differentials among the products.

Recall that for any demand instance $\boldsymbol{d} \in \mathbb{R}^n_+$, the maximum profit achieved by flexibility design $\mathcal{A}$, $g(\mathcal{A}, \mathbf{d})$, can be computed by solving (1). By strong duality theorem, we have

$$
\begin{aligned}
g(\mathcal{A}, \mathbf{d}) = \min_{y,z} & \sum_{i=1}^{n} d_i y_i + \sum_{j=1}^{m} C_j z_j \qquad (3) \\
\text{s.t.} \quad & y_i + z_j \geq p_i, \quad \forall (a_j, b_i) \in \mathcal{A}, \\
& y_i \geq 0, \quad \forall 1 \leq i \leq n, \\
& z_j \geq 0, \quad \forall 1 \leq j \leq m
\end{aligned}
$$

We first present the following lemma that provides a characterization of the optimal solutions to the dual problem (3), which will become useful in defining the dual margin group index later. All the proofs in this paper are relegated to Appendix A.

Lemma 1. *Suppose (y^*, z^*) is a basic feasible solution to the dual problem (3). For any product $i \in \{1, ..., n\}$, one of the following three cases must be true: (i) $y_i^* = p_i$, (ii) $y_i^* = 0$, or (iii) $y_i^* = p_i - p_{i'}$ for some $p_{i'} < p_i$.*

In view of Lemma 1, there must exist an optimal solution (y^*, z^*) to the dual problem (3) where y_i^* can only take one of the values from the set $\{p_i\} \cup \{0\} \cup \{p_i - p_{i'} \mid 1 \leq i' \leq n \text{ such that } p_{i'} < p_i\}$ for each product i. Intuitively, the dual variable y_i^* represents the shadow price associated with the product i demand d_i. In the case of constrained capacity, the "marginal value" brought by an additional unit of product i demand can be either zero (i.e., this additional unit of product i demand cannot be fulfilled due to insufficient capacity), or the difference between its own margin and that of a less profitable product (in which case an extra unit of product i demand is satisfied at the expense of giving up the sale of some lower profitable product). When the capacity is unconstrained, this extra unit of product i demand is fulfilled and brings an additional profit margin p_i.

Next, we shall use the above observation to define the DMGI to facilitate the evaluation and comparison of flexibility designs. For each $1 \le t \le T$, define $\beta_{tr} := P_t - P_r$ for each $0 \le r \le t$, where $P_0 \equiv 0$. For a fixed $\boldsymbol{\mu}$, we define the *Dual Margin Group Index* (DMGI) at parameter set $K = \{k_{tr} \in \mathbb{Z}^+ : t = 1, ..., T,\, r = 0, ..., t-1\}$ for flexibility design $\mathcal{A}$, denoted by $\delta_{\boldsymbol{\mu}}^K(\mathcal{A})$, as the objective value of the following integer program:

$$\delta_{\boldsymbol{\mu}}^K(\mathcal{A}) := \min_{\boldsymbol{w},\boldsymbol{y},\boldsymbol{z}} \quad \left\{ \sum_{i=1}^n \mu_i y_i + \sum_{j=1}^m C_j z_j \right\} \tag{4}$$

$$\begin{aligned}
\text{s.t.} \quad & y_i + z_j \ge p_i, \quad \forall (a_j, b_i) \in \mathcal{A} \\
& y_i = \sum_{r=0}^t \beta_{tr} w_{itr}, \quad \forall\, b_i \in S_t,\, \forall t = 1, ..., T \\
& \sum_{b_i \in S_t} w_{itr} = k_{tr}, \quad \forall\, t = 1, ..., T,\, \forall\, r = 0, ..., t-1 \\
& \sum_{r=0}^t w_{itr} = 1, \quad \forall\, b_i \in S_t,\, \forall t = 1, ..., T \\
& \sum_{r=0}^t w_{itr} = 0, \quad \forall\, b_i \notin S_t,\, \forall t = 1, ..., T \\
& w_{itr} \in \{0, 1\} \quad \forall\, b_i \in S_t,\, \forall\, r = 0, ..., t,\, \forall t = 1, ..., T
\end{aligned}$$

A few remarks are in order. By Lemma 1, we have $y_i \in \{\beta_{tr} \,|\, 0 \le r \le t\}$ for any product i with margin P_t if y_i comes from a basic feasible solution to the dual problem (3). In the above definition, the binary variable $w_{itr} = 1$ if y_i takes value β_{tr} for product $b_i \in S_t$. The parameter k_{tr} prescribes that there are exactly k_{tr} products with margin P_t whose y_i is equal to β_{tr}, i.e., $|\{b_i \in S_t : y_i = \beta_{tr}\}| = k_{tr}$. Notice that $\beta_{tt} = 0$ for all $t = 1, ..., T$. It then follows that the cardinality of the set $\{b_i \in S_t : y_i = \beta_{tt}\}$ does not affect objective value (4), and therefore the parameters k_{tr} are only defined for $r = 0, ..., t-1$. Intuitively, $\delta_{\boldsymbol{\mu}}^K(\mathcal{A})$ represents the minimum requirement of the mean demand and plant capacities to ensure dual feasibility, given that the values of y_i for each product group satisfy the cardinality constraint specified by parameter set K. It is appropriate to remark that although computing $\delta_{\boldsymbol{\mu}}^K(\mathcal{A})$ is NP-hard (cf. Lemma 2 in Simchi-Levi and Wei (2015) for a special case of our model), the integer program formulation (4) in general has a small size and can be solved efficiently using commercial optimization solvers such as Gurobi. With (4), we show in the next subsection that the DMGI is a useful tool for the evaluation and comparison of the worst-case performance of different flexibility designs.

3.2. Worst-Case Measures with Part-wise Independently Symmetric Perturbation Uncertainty Sets

For any given flexibility design $\mathcal{A}$, we study its robust measures under part-wise independently symmetric perturbation uncertainty sets in this subsection. Consider a part-wise independently symmetric perturbation uncertainty set U for some fixed $\boldsymbol{\mu}$. For each index set $s_t \subset [n]$, let $\boldsymbol{d}_{s_t} = \{d_i \,|\, b_i \in S_t\}$ and $\boldsymbol{\mu}_{s_t} = \{\mu_i \,|\, b_i \in S_t\}$. For a given parameter set $K = \{k_{tr} \in \mathbb{Z}^+ : t = 1, ..., T, r = 0, ..., t-1\}$ where $0 \le \sum_{r=0}^{t-1} k_{tr} \le |S_t|$, let $H_{tr} := \sum_{j=0}^{r} k_{tj}$ for all $1 \le t \le T$, $0 \le r \le t-1$. Let Q denote the set of feasible non-negative integer parameter sets $K = \{k_{tr} \in \mathbb{Z}^+ : t = 1, ..., T, r = 0, ..., t-1\}$ such that $0 \le \sum_{r=0}^{t-1} k_{tr} \le |S_t|$. For any robust measure $R^f \in \Gamma$, we show in the following result that there exists an explicit representation of the worst-case performance $R^f(\mathcal{A}, U)$ for any design $\mathcal{A}$ and part-wise independently symmetric perturbation uncertainty set U.

THEOREM 1. *Let f be a deterministic measure function that is monotonic in profit under fixed total demand in each product category $\sum \boldsymbol{d}_{s_t}$ and separable in $\sum \boldsymbol{d}_{s_t}$ for each $1 \le t \le T$. Let $h(\cdot)$ be the function such that $h(x, y_1, ..., y_T)$ is strictly increasing in x for fixed $y_1, ..., y_T$, and $f(\mathcal{A}, \boldsymbol{d}) = h(g(\mathcal{A}, \boldsymbol{d}), \sum \boldsymbol{d}_{s_1}, ..., \sum \boldsymbol{d}_{s_T})$. Then for any uncertainty set U that is part-wise independently symmetric around $\boldsymbol{\mu}$, we have*

$$R^f(\mathcal{A}, U) = \min_{K \in Q, \boldsymbol{\epsilon} \in E} h\left(\delta_{\boldsymbol{\mu}}^K(\mathcal{A}) + \sum_{t=1}^{T}\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}} \epsilon_{s_t,i}\right)\right), \sum_{i \in s_1}(\mu_i + \epsilon_i), ..., \sum_{i \in s_T}(\mu_i + \epsilon_i)\right).$$

Theorem 1 shows that for any robust measure $R \in \Gamma$ and part-wise independently symmetric perturbation uncertainty set U, the worst-case performance of design $\mathcal{A}$ under robust measure R

can be completely determined by the values of $\delta_{\boldsymbol{\mu}}^K(\mathcal{A})$ without any additional information of $\mathcal{A}$. Moreover, Theorem 1 implies a partial order of flexibility designs under any robust measure in Γ.

Theorem 2. *Fix a robust measure $R \in \Gamma$. Then the following statements hold:*

(a) If $\delta^K(\mathcal{A}_1, U) \geq \delta^K(\mathcal{A}_2, U)$ for all $K \in Q$, then $R(\mathcal{A}_1, U) \geq R(\mathcal{A}_2, U)$ for any part-wise independently symmetric perturbation uncertainty set U.

(b) If $R(\mathcal{A}_1, U) \geq R(\mathcal{A}_2, U)$ for any part-wise independently symmetric perturbation uncertainty set U, then $\delta^K(\mathcal{A}_1, U) \geq \delta^K(\mathcal{A}_2, U)$ for all $K \in Q$ such that $k_{tr} = 0$ for all $r > 0$.

Theorem 2(a) provides a sufficient condition for one flexibility design to be more robust (i.e., with a better worst-case performance) than another. As we shall show later, this sufficient condition allows us to compare the worst-case performance of different designs, and is used to establish the optimality of a special class of long chain, the so-called *alternate long chain design*, among all the long chain designs with equal number of high margin and low margin products. However, we have been unable to show the necessity of this condition, and Theorem 2(b) presents a necessary condition when $R(\mathcal{A}_1, U) \geq R(\mathcal{A}_2, U)$ for any part-wise independently symmetric perturbation uncertainty set U. By Theorem 2(b), if one design is more robust than another, then the DMGI value of the former must be no less than that of the latter under certain parameter sets. Later in Section 4, we will use this necessary condition to show the suboptimality of the alternate long chain among the class of 2-flexibility designs with equal number of high profit and low profit products.

4. Worst-Case Performance of the Alternate Long Chain Design

One class of flexibility designs that has been extensively studied in the operations literature is the long chain $\mathcal{L}$, which has been shown to be extremely effective in mitigating supply-demand mismatch where the evaluation metric is based on the total sales in units. In this section, we aim to study the worst-case effectiveness of the long chain design when there are two product categories with respective profit margins P_H and P_L. Let $S_H = \{b_i \in B \mid p_i = P_H\}$ and $S_L = \{b_i \in B \mid p_i = P_L\}$ denote the set of *high profit products* and the set of *low profit products*, respectively.

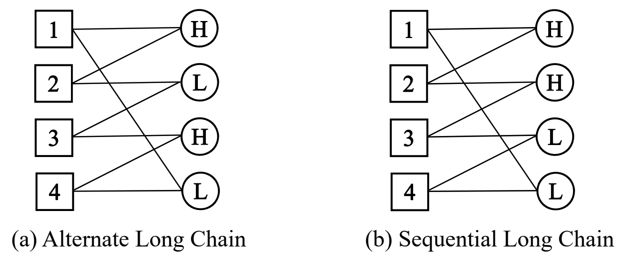
Throughout this section, we consider a balanced system where the numbers of the plants and products are both equal to n. As is typical in the analysis of the long chain (e.g., Simchi-Levi and Wei (2012, 2015), Wang and Zhang 2015), we assume that all the plants have identical capacities and all the products have the same expected demand, i.e., $C_i = C_j$ and $\mu_i = \mu_j$ for all $1 \le i, j \le n$. Without loss of generality, we assume unit capacities $C_j = 1$ for all $1 \le j \le n$.

Let $K = \{k_{10}, k_{20}, k_{21}\}$ be a given parameter set, where $k_{10} = \{b_i \in S_L \,|\, y_i = P_L\}$, $k_{20} = \{b_i \in S_H \,|\, y_i = P_H\}$, $k_{21} = \{b_i \in S_H \,|\, y_i = P_H - P_L\}$, and y_i is the variable corresponding to each product in the DMGI definition (4). When $\mu_i = \mu_j$ for all $1 \le i, j \le n$, the term $\sum_{i=1}^{n} \mu_i y_i$ in the objective function (4) is a constant for any fixed parameter set K. Therefore, we may omit this term without loss of optimality and the DMGI for a homogeneous design $\mathcal{A}$, denoted by $\delta^K(\mathcal{A})$, is equivalent to the optimal objective function of the following problem:

$$
\begin{aligned}
\delta^K(\mathcal{A}) := \min_{\boldsymbol{w},\boldsymbol{y},\boldsymbol{z}} \quad & \sum_{j=1}^{m} z_j && (5)\\
\text{s.t.} \quad & y_i + z_j \ge p_i, \quad \forall (a_j, b_i) \in \mathcal{A} \\
& y_i = \sum_{r=0}^{t} \beta_{tr} w_{itr}, \quad \forall\, b_i \in S_t,\, \forall t = 1, ..., T \\
& \sum_{b_i \in S_t} w_{itr} = k_{tr}, \quad \forall\, t = 1, ..., T,\, \forall\, r = 0, ..., t-1 \\
& \sum_{r=0}^{t} w_{itr} = 1, \quad \forall\, b_i \in S_t,\, \forall t = 1, ..., T \\
& \sum_{r=0}^{t} w_{itr} = 0, \quad \forall\, b_i \notin S_t,\, \forall t = 1, ..., T \\
& w_{itr} \in \{0, 1\} \quad \forall\, b_i \in S_t,\, \forall\, r = 0, ..., t,\, \forall t = 1, ..., T
\end{aligned}
$$

Consider the class of long chain designs with equal number of high profit products and low profit products. Among this class of long chain designs with $|S_H| = |S_L| = n/2$, we are particularly interested in a special design where the high profit products and low profit products are assigned to the plants in an alternate manner, which we call the *alternate long chain design* and is denoted as $\mathcal{AL}$. Figure 2(a) illustrates an alternate long chain with $n = 4$, where each circle represents a product node and the high profit products (H) and low profit products (L) are spread out evenly in an alternate manner. In the remainder of this section, we apply the results from the previous sections to analyze the worst-case effectiveness of the alternate long chain design.

Figure 2 Alternate Long Chain and Sequential Long Chain with $n = 4$



4.1. Optimality of the Alternate Long Chain Design

In this subsection, we first examine the performance of the alternate long chain among all the long chain designs with equal number of high profit products and low profit products $\mathcal{L}_{HL}$. Our next result demonstrates that the alternative long chain design has the best worst-case performance among $\mathcal{L}_{HL}$ under any robust measure R that lies in Γ.

THEOREM 3. *Let $\mathcal{L}_{HL}$ be a long chain design with equal number of high profit products and low profit products. Then, for any part-wise independently symmetric uncertainty set U and any robust measure $R \in \Gamma$, we have $R(\mathcal{AL}, U) \geq R(\mathcal{L}_{HL}, U)$.*

In view of Theorem 3, the alternate long chain is optimal among all the long chain designs with equal number of high profit products and low profit products. Intuitively, among all the long chain designs with equal number of high profit and low profit products, the alternate long chain has the best potential to fulfill the demand from high profit products compared with designs where multiple high profit products share a common plant. One implication of the above optimality result is that, within the class of long chain production systems, it is most beneficial to evenly "spread out" the products from the two product groups in an alternate manner and to equip each plant to be capable of producing both high profit and low profit products. As a result, the high profit products never have to directly compete for shared resources. A natural generalization of Theorem

3 is to investigate the optimality of the alternate long chain design among the class of 2-flexibility designs where each product is connected to two plants. However, as we show in Section 4.2, the optimality result in Theorem 3 does not extend to the more general class of 2-flexibility designs in the presence of margin differentials among the products.

We next give a sketch of the proof of Theorem 3 with the complete proof relegated to Appendix A. In Lemma 2, we provide the exact characterization of the DMGI value for the alternate long chain design $\delta^K(\mathcal{AL})$ for any given parameter set K. We then show in Proposition 1 that the alternate long chain design has the largest DMGI for any $K \in Q$ among all long chain designs with equal number of high profit and low profit products, which immediately implies Theorem 3 by the sufficient condition in Theorem 2(a).

LEMMA 2. *For a given parameter set $K = \{k_{10}, k_{20}, k_{21}\}$, the dual margin group index of the alternate long chain design $\mathcal{AL}$ with $n/2$ high profit products and $n/2$ low profit products is given by*

$$\delta^K(\mathcal{AL}) = \begin{cases} 2k_{22}P_H + 2k_{21}P_L, & \text{if } k_{20} = 0 \\ 2k_{22}P_H + (n - 2k_{22} - \min(k_{10}, k_{20}+1) - \min(k_{10}, k_{20}-1))P_L, & \text{if } k_{20} = 1, \dots, \frac{n}{2}-1, \\ 2k_{11}P_L, & \text{if } k_{20} = \frac{n}{2}, \end{cases} \tag{6}$$

where $k_{22} = n/2 - k_{20} - k_{21}$ and $k_{11} = n/2 - k_{10}$.

PROPOSITION 1. *Let $\mathcal{L}_{HL}$ be a long chain design with equal number of high profit products and low profit products. Then, $\delta^K(\mathcal{AL}) \geq \delta^K(\mathcal{L}_{HL})$ for any parameter set $K \in Q$.*

Theorem 3 establishes the optimality of the alternate long chain among all the long chain designs with equal number of high profit and low profit products in a balanced production system where all the plants have identical capacities and all the products have the same expected demand. Intuitively, compared with other long chain designs in which multiple high profit products may share the same plant, the alternate long chain can achieve superior robust performance since high profit products will never directly “compete” with each other for the limited capacities.

We have conducted numerical analysis to further illustrate the intuition for the superiority of the alternate long chain. We briefly summarize the numerical studies below, and we refer the readers to Section 4.2 in the unabridged version Wang et al. (2020) for the details. Our numerical studies compare the performance of the alternate long chain and the sequential long chain, where in a *sequential long chain*, all the high profit products are adjacent to each other and all the low profit products are grouped together in a sequence (see Figure 2(b) for an illustration of a sequential long chain with $n=4$). We evaluate the profits achieved by the two long chain designs under randomly generated demand instances and compute the average ratio and worst-case ratio of their profits relative to full flexibility under various parameter settings. Our simulation results suggest that the alternate long chain considerably outperforms the sequential long chain under both average and worst-case performance measures. We have extended our numerical analysis to compare the performance of different long chain designs with three and four different profit levels, and further generalized to unbalanced systems in which the numbers of the products and plants are not equal. Our numerical results demonstrate that the insights obtained from the optimality of the alternate long chain that products with different margins should be "spread out" are quite robust, which also apply to systems with more than two profit levels and unbalanced production systems.

4.2. Alternate Long Chain vs. Disjoint Long Chain

As shown in Theorem 3, the alternate long chain is optimal among all the long chain designs with equal number of high profit and low profit products. A natural generalization to Theorem 3 is to investigate the optimality of the alternate long chain among the class of 2-flexibility designs with equal number of high profit and low profit products, where in a *2-flexibility design*, each product is connected to exactly two plants. However, as we shall show below, this generalization does not hold in general. In particular, we compare the performance of the alternate long chain with a specific class of 2-flexibility designs, the *disjoint long chain*, and demonstrate that the alternate long chain could be either superior or inferior to the disjoint long chain. A disjoint long chain $\mathcal{DL}$ with n products is composed of two separate long chain components with $n/2$ products, where the products that belong to the same component have the same profit (a $\mathcal{DL}$ with $n = 8$ is illustrated in Wang et al. (2020) Appendix F). Our next result shows that there is no dominance relationship between the disjoint long chain and the alternate long chain under any robust measure in Γ.

Proposition 2. *Fix a robust measure $R \in \Gamma$. Then there exist some part-wise independently symmetric uncertainty sets U and U' such that $R(\mathcal{AL}, U) \geq R(\mathcal{DL}, U)$ and $R(\mathcal{AL}, U') \leq R(\mathcal{DL}, U')$.*

In view of Proposition 2, the disjoint long chain may achieve a better worst-case performance than the alternate long chain, and therefore the alternate long chain design is no longer optimal among the class of 2-flexibility designs. It is worth noticing that the disjoint long chain differs from the sequential long chain in that the high profit products and low profit products are completely disconnected in the disjoint long chain. Therefore, the high profit products have full access to half of the total plant capacities that will not be shared by any low profit product either directly or indirectly. In contrast, the alternate long chain is a *connected* design and hence high profit products may need to indirectly compete for capacities with low profit products and result in an inferior worst-case performance than the disjoint long chain. An example that shows the suboptimality of the alternate long chain is presented in Wang et al. (2020) Appendix D.

5. Generating Effective Flexibility Design under Profit Maximization

In this section, we aim to develop a heuristic by deploying the notion of DMGI for constructing effective flexibility designs when there exist margin differentials among the products. The heuristic we propose start with an initial base flexibility design $\mathcal{A}$ (which could be an empty system without any arc), and iteratively add arcs to improve $\mathcal{A}$. In what follows, we first describe the high level idea of a single iteration and then present the formal description of our proposed DMGI-Heuristic in Algorithm 1, followed by numerical analysis to evaluate the effectiveness of our heuristic.

Recall that Q denotes the set of feasible non-negative integer parameter sets $K = \{k_{tr} \in \mathbb{Z}^+ : t = 1, ..., T, r = 0, ..., t-1\}$ such that $0 \leq \sum_{r=0}^{t-1} k_{tr} \leq |S_t|$, where for each $0 \leq r \leq t$, the integer k_{tr} prescribes that there are exactly k_{tr} products with profit margin P_t whose y_i value in the definition of $\delta^K_{\boldsymbol{\mu}}(\mathcal{A})$ is equal to $P_t - P_r$. In view of Theorem 2(a), it is desired to find designs with large

DMGI values $\delta_{\boldsymbol{\mu}}^{K}(\mathcal{A})$ for all $K \in Q$ in order to construct flexibility designs with effective robust performance. Therefore, the main idea of our algorithm is to add an arc $(a_j, b_i) \notin \mathcal{A}$ to $\mathcal{A}$ in order to increase the values of DMGI at each iteration as much as possible.

We next introduce some necessary notations to formally present our heuristic. Recall that DMGI is defined under a given parameter set $K = \{k_{tr} \in \mathbb{Z}^+ : t = 1, ..., T, r = 0, ..., t-1\}$, where k_{tr} represents the number of products with margin P_t whose y_i value is equal to $P_t - P_r$. For notation convenience, let k_{tt} denote the number of products with margin P_t whose y_i value is equal to 0, i.e., $k_{tt} = |\{b_i \in S_t : y_i = 0\}|$. Notice that the value of k_{tt} is uniquely determined for each t under a fixed parameter set K. Instead of considering all possible $K \in Q$, we focus on the K that has the worst DMGI value within the same group $Q_{(q_1,...,q_T)} \subset Q$ parameterized by integers $\{q_1, ..., q_T\}$, where

$$Q_{(q_1,...,q_T)} := \left\{ K \in Q \,\Big|\, \sum_{t=r}^{T} k_{tr} = q_r, \forall 1 \le r \le T \right\},$$

and parameter q_r prescribes the number of products whose margin P_t is at least P_r and its y_i value in the definition of $\delta_{\boldsymbol{\mu}}^{K}(\mathcal{A})$ is equal to $P_t - P_r$. For each $\boldsymbol{q} := (q_1, ..., q_T)$ where $0 \le q_r \le \sum_{t=r}^{T} |S_t|$ for each $1 \le r \le T$, we solve for the worst DMGI value among all $K \in Q_{\boldsymbol{q}}$:

$$K^{\boldsymbol{q}} := \operatorname*{arg\,min}_{K \in Q_{\boldsymbol{q}}} \delta_{\boldsymbol{\mu}}^{K}(\mathcal{A}),$$

and obtain the corresponding optimal solution $\boldsymbol{y}^{K^{\boldsymbol{q}}}$ and $\boldsymbol{z}^{K^{\boldsymbol{q}}}$. Notice that the objective function (4) in the definition of $\delta_{\boldsymbol{\mu}}^{K}(\mathcal{A})$ increases with the value of y_i and z_j. In view of this, we define pair (i, j) as a *bottleneck* under $K^{\boldsymbol{q}}$ if both $y_i^{K^{\boldsymbol{q}}} = 0$ and $z_j^{K^{\boldsymbol{q}}} = 0$. Intuitively, as y_i and z_j are dual variables associated with the demand and capacity constraints in the max-flow problem (1), in view of the complementary slackness condition, a bottleneck prevents the system from using the available plant capacities to satisfy the excess demand. In each iteration, our proposed heuristic computes $K^{\boldsymbol{q}}$ for all $\boldsymbol{q}$ and then adds an arc that can reduce the largest number of bottlenecks, which would be helpful in obtaining a larger value of the dual margin group index. Our numerical studies demonstrate that the performance of the sparse designs generated by our DMGI-Heuristic compares very well with that of full flexibility, and the value of limited flexibility becomes more

Algorithm 1 DMGI-Heuristic

1: Input: An initial design $\mathcal{A}$ in an m plants n products system, and a budget of extra B arcs.

2: **for** $b=1,...,B$ **do**

3: Compute $K^{\boldsymbol{q}}$ and its corresponding optimal solution $(\boldsymbol{y}^{K^{\boldsymbol{q}}}, \boldsymbol{z}^{K^{\boldsymbol{q}}})$ for each $\boldsymbol{q} := (q_1,...,q_T)$ such that $0 \le q_r \le \sum_{t=r}^{T} |S_t|$ for each $1 \le r \le T$.

4: For each $1 \le i \le n, 1 \le j \le m$ such that $(a_j, b_i) \notin \mathcal{A}$, compute

$$W(i,j) = \sum_{\boldsymbol{q}} \mathbb{1}\{y_i^{K^{\boldsymbol{q}}} = 0\} \times \mathbb{1}\{z_j^{K^{\boldsymbol{q}}} = 0\}.$$

5: Find arc (a_{j^*}, b_{i^*}) such that $W(i^*, j^*) = \max_{1\le i\le n, 1\le j\le m} W(i,j)$. When there is a tie, we uniformly select an arc with the maximum $W(i,j)$.

6: Add arc (a_{j^*}, b_{i^*}) to $\mathcal{A}$.

7: **end for**

8: Return $\mathcal{A}$

significant when the demand has a lower volatility and the margin difference is smaller. We refer the readers to Wang et al. (2020) Section 5 for details.

Our DMGI-Heuristic is developed based on a robust optimization based approach. Another approach to designing flexibility structures in the case of known demand distribution is to solve a two-stage integer stochastic programming problem, and we provide some discussions on our choice of robust optimization approach over stochastic optimization approach in Wang et al. (2020) Appendix G. We next present numerical studies that compare the performance achieved by the flexibility designs generated by our DMGI-Heuristic and designs obtained by stochastic optimization based approaches. In particular, we compared our algorithm with Sample Average Approximation (SAA) method and Feng et al. (2017), and the numerical results suggest that our robust optimization approach outperform the other two stochastic optimization based approaches, especially when there exist demand distribution misspecification. More specifically, we consider a production system with 14 products and 7 plants. Half of the products have profit margin equal to \$2, and their demands are independent and follow Normal$(1, 0.4^2)$ truncated at zero. The other half of the products have margin equal to \$1, and their demands are independent and follow Normal$(2, 0.8^2)$

truncated at zero. All the plants have the same capacity of 3 units. In Table 1, we report the numerical results that compare the performance of our DMGI-Heuristic, the stochastic optimization based heuristic proposed by Feng et al. (2017) (the MDEP-Heuristic), and the SAA method, where we start with an empty system and add a total number of 25 arcs to the system.

In our numerical experiments, we randomly drew 2000 demand samples to generate a design using the MDEP-Heuristic by Feng et al. (2017), and used 200 demand samples to generate a flexibility structure using the SAA method. Then we compare the above two designs with the one generated by DMGI-Heuristic under three scenarios. In Scenario 1, the actual demand distribution is the same as the original distribution (i.e., the demands of half of the products are independent and follow Normal$(1, 0.4^2)$ truncated at zero, and the demands of the other half are independent and follow Normal$(2, 0.8^2)$ truncated at zero). In Scenario 2, there exist demand distribution misspecification. For the actual demand, each product still follows their original distribution as aforementioned, but the demands within each class are positively correlated with correlation coefficient equal to 0.4. Scenario 3 is similar to Scenario 2, but the correlation coefficient is equal to 0.8. For all three scenarios, we have conducted 100 simulation runs, each with 10000 demand instances. The mean (in bold font) and standard deviation (in parenthesis) of the average ratios and the worst-case ratios of the profits achieved by the three heuristics relative to full flexibility are presented in Table 1. We observe that our DMGI-Heuristic slightly outperform MDEP-Heuristic and SAA when the actual demand follows the original distribution. Moreover, the advantage of our algorithm compared with the other two benchmarks becomes larger as the correlation coefficient increases.

Since the majority of the heuristics for generating effective sparse flexibility designs proposed in the literature aim to achieve the maximum sales in units without taking margin differentials into account, we have also conducted another set of numerical studies in the setting where $P_H = P_L = 1$. Our benchmarking algorithms include the Plant Cover Index (PCI) Heuristic by Simchi-Levi and Wei (2015) and the Expander heuristic developed in Chou et al. (2011), which have been shown to outperform many existing algorithms in the literature such as the algorithm proposed by Hopp et al. (2004) as well as the design with the highest expected sales among 50 randomly generated designs

Table 1 Average Ratio and Worst-case Ratio between the Profits of the Sparse Designs with 25 arcs by Various Heuristics relative to Full Flexibility under Independent (Scenario 1), Positively Correlated with Correlation Coefficient 0.4 (Scenario 2) and Positively Correlated with Correlation Coefficient 0.8 (Scenario 3) Demand

	Average Ratio			Worst-case Ratio		
	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
DMGI-Heuristic	**0.9995** (0.0000)	**1.0000** (0.0000)	**1.0000** (0.0000)	**0.9304** (0.0093)	**0.9629** (0.0075)	**0.9981** (0.0036)
MDEP-Heuristic	**0.9993** (0.0000)	**0.9999** (0.0000)	**1.0000** (0.0000)	**0.9251** (0.0109)	**0.9573** (0.0086)	**0.9594** (0.0115)
SAA	**0.9993** (0.0000)	**0.9999** (0.0000)	**1.0000** (0.0000)	**0.9209** (0.0118)	**0.9562** (0.0096)	**0.9576** (0.0129)

(e.g. Simchi-Levi and Wei 2015). Here we would like to point out that our DMGI-Heuristic is closely related to the PCI-Heuristic, and we provide a detailed discussion on how the DMGI-Heuristic differs from the PCI-Heuristic in the case where all the products have identical profit margins in Wang et al. (2020) Appendix E. We have also considered the well-known k-chain structures, which are known to be effective in balanced and symmetric systems (e.g. Chou et al. 2014).

Table 2 summarizes the average ratio and the worst-case ratio of the profits achieved by various sparse designs in a 10 by 10 production system with 25 arcs and 30 arcs (generated starting from an empty graph) relative to that of full flexibility when all the products have identical profit margins. The corresponding flexibility designs generated by the DMGI-Heuristic, the PCI-Heuristic, and the Expander heuristic are presented in Wang et al. (2020) Appendix F. In addition to the PCI-Heuristic and the Expander heuristic, we have also compared our algorithm with the chaining structures. In the case of a budget with 25 arcs, we consider a so-called 2.5-chain, which is a long chain (or, 2-chain) with 5 additional arcs $\{(a_6, b_1), (a_7, b_2), (a_8, b_3), (a_9, b_4), (a_{10}, b_5)\}$ that are evenly distributed among the product-plant pairs (see Figure 7 in Appendix F for an illustration). From Table 2, we observe that our DMGI-Heuristic has comparable performance with PCI-Heuristic in terms of average ratios, and achieves a slightly better worst-case ratio than that by PCI-Heuristic. For the Expander algorithm and the chaining structures, our DMGI-Heuristic consistently achieves a better performance from both the average and the worst-case perspective, and the improvement in the worst-case ratio becomes more significant as the demand becomes more volatile. These observations suggest that our proposed DMGI-Heuristic can be a favorable algorithm for generating effective sparse designs even when the products have identical profit margins. We have also conducted another numerical study for a larger system with 40 plants and 40 products, and the results (summarized in Appendix F) suggest similar insights as that observed from Table 2.

Table 2 Average Ratio and Worst-case Ratio of the Profits achieved by Various Heuristics with the same number of arcs relative to Full Flexibility in a 10 by 10 Production System: $P_H = P_L = 1$ and $D \sim N(1, \sigma^2)$ truncated at zero

25 arcs		DMGI-Heuristic	PCI-Heuristic	Expander	2.5-Chain
$\sigma = 0.4$	Average Ratio	**0.9984** (0.0001)	**0.9983** (0.0001)	**0.9560** (0.0004)	**0.9973** (0.0001)
	Worst-case Ratio	**0.8714** (0.0196)	**0.8833** (0.0150)	**0.7572** (0.0204)	**0.8592** (0.0165)
$\sigma = 0.8$	Average Ratio	**0.9754** (0.0004)	**0.9730** (0.0004)	**0.9164** (0.0006)	**0.9701** (0.0005)
	Worst-case Ratio	**0.6800** (0.0291)	**0.6973** (0.0297)	**0.5663** (0.0305)	**0.6681** (0.0269)
30 arcs		DMGI-Heuristic	PCI-Heuristic	Expander	3-Chain
$\sigma = 0.4$	Average Ratio	**1.0000** (0.0000)	**0.9998** (0.0000)	**0.9708** (0.0003)	**0.9999** (0.0000)
	Worst-case Ratio	**0.9506** (0.0127)	**0.9313** (0.0129)	**0.7716** (0.0163)	**0.9369** (0.0182)
$\sigma = 0.8$	Average Ratio	**0.9936** (0.0002)	**0.9923** (0.0002)	**0.9459** (0.0006)	**0.9913** (0.0002)
	Worst-case Ratio	**0.7778** (0.0241)	**0.7727** (0.0291)	**0.5926** (0.0281)	**0.7416** (0.0237)

Up to this point, our numerical studies have focused on balanced systems. We have conducted an additional set of numerical studies on systems with asymmetric capacities and non-*i.i.d.* demand distributions. In particular, we compare our DMGI-Heuristic with three known algorithms, the Thresholded Probabilistic Construction (TPC) heuristic by Chen et al. (2019), the Generalized Chaining Gap (GCG) based heuristic developed by Shi et al. (2019) and the Expander heutistic by Chou et al. (2011) with details given in Wang et al. (2020) Section 5. Our results suggest that our DMGI-Heuristic and the Expander heuristic can achieve a better performance than the other two approaches in an unbalanced and asymmetric system, especially when there exists demand distribution misspecification, which can provide a higher level of robustness of the performance.

6. Concluding Remarks

In this paper, we have studied the worst-case performance of flexibility designs when there exist margin differentials among the products and the evaluation metric is based on maximizing total profits rather than sales in units. Inspired by the structural property of the solutions to the dual formulation of the profit maximization problem, we introduce the Dual Margin Group Index (DMGI) and prove that a general class of worst-case performance measures can be expressed as functions of the DMGIs under part-wise independently symmetric uncertainty sets. One implication of this result is that the set of all DMGIs provides a sufficient statistic to compute the worst-case performance of any flexibility design without any additional information of the design, which, in turn, leads to a partial ordering that enables us to compare the performance of different flexibility designs. By applying the above results, we establish the optimality of the alternate long chain among all long chain designs where there are two product categories and the number of high profit products is equal to that of low profit products. Motivated by our theoretical results, we propose a heuristic based on DMGI for generating effective flexibility designs when the products exhibit margin differentials. Our computational study suggests that the sparse design generated by DMGI-Heuristic captures most of the benefit of full flexibility from the expected profit point of view, and has satisfactory performance in the worst case for small to moderate level of demand uncertainty.

There are several ways to extend our research. First, in Theorem 2(a) we identify a sufficient condition $\delta^K(\mathcal{A}_1, U) \geq \delta^K(\mathcal{A}_2, U)$ for all $K \in Q$ for a design $\mathcal{A}_1$ to outperform another design $\mathcal{A}_2$ for any part-wise independently symmetric perturbation uncertainty set U under any robust measure $R \in \Gamma$. This sufficient condition is stronger than the necessary condition in Theorem 2(b), and an interesting open question is whether we can close the gap to get a condition that is both sufficient and necessary. Second, the comparison of different flexibility designs' worst-case performance and the DMGI-heuristic require computing $\delta^K_{\boldsymbol{\mu}}(\mathcal{A})$ for all $K \in Q$. For small to medium sized systems, our numerical experiments suggest that the optimization problem for computing $\delta^K_{\boldsymbol{\mu}}(\mathcal{A})$ can be solved efficiently. However, for large size systems, it is important to develop efficient methods for the computation of $\delta_{\boldsymbol{\mu}}^{K}(\mathcal{A})$ to make DMGIs applicable in more practical settings. Finally, our current model only allows product-dependent profit margins. It would be nice to develop models and solution approaches that allow the profit margins to be plant-product pair specific.

References

Asadpour, Arash, Xuan Wang, Jiawei Zhang. 2020. Online resource allocation with limited flexibility. *Management Science* **66**(2) 642–666.

Bassamboo, Achal, Ramandeep S. Randhawa, Jan A. Van Mieghem. 2010. Optimal flexibility configurations in newsvendor networks: Going beyond chaining and pairing. *Management Science* **56**(8) 1285–1303.

Bertsimas, Dimitris, Vineet Goyal, Xu Andy Sun. 2011. A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Mathematics of Operations Research* **36**(1) 24–54.

Buzacott, John A., Marvin Mandelbaum. 2008. Flexibility in manufacturing and services: achievements, insights and challenges. *Flexible Services and Manufacturing Journal* **20**(1-2) 13–58.

Chen, Xi, Tengyu Ma, Jiawei Zhang, Yuan Zhou. 2019. Optimal design of process flexibility for general production systems. *Operations Research* **67**(2) 516–531.

Chen, Xi, Jiawei Zhang, Yuan Zhou. 2015. Optimal sparse designs for process flexibility via probabilistic expanders. *Operations Research* **63**(5) 1159–1176.

Chevrolet. 2010. Chevrolet equinox production begins at oshawa assembly, increasing equinox and gmc terrain capacity. Available at `https://media.gm.com/media/us/en/chevrolet/vehicles/Equinox/2011.detail.html/content/Pages/news/us/en/2010/Sept/0920_Equinox.html`.

Chou, Mabel C., Geoffrey A. Chua, Chung-Piaw Teo. 2010a. On range and response: Dimensions of process flexibility. *European Journal of Operational Research* **207**(2) 711–724.

Chou, Mabel C., Geoffrey A. Chua, Chung-Piaw Teo, Huan Zheng. 2010b. Design for process flexibility: Efficiency of the long chain and sparse structure. *Operations Research* **58**(1) 43–58.

Chou, Mabel C., Geoffrey A. Chua, Chung-Piaw Teo, Huan Zheng. 2011. Process flexibility revisited: The graph expander and its applications. *Operations Research* **59**(5) 1090–1105.

Chou, Mabel C., Geoffrey A. Chua, Huan Zheng. 2014. On the performance of sparse process structures in partial postponement production systems. *Operations Research* **62**(2) 348–365.

Chou, Mabel C., Chung-piaw Teo, Huan Zheng. 2008. Process flexibility: design, evaluation, and applications. *Flexible Services and Manufacturing Journal* **20**(1-2) 59–94.

Désir, Antoine, Vineet Goyal, Yehua Wei, Jiawei Zhang. 2016. Sparse process flexibility designs: Is long chain really optimal? *Operations Research* **64**(2) 416–431.

Feng, Wancheng, Chen Wang, Zuo-Jun Max Shen. 2017. Process flexibility design in heterogeneous and unbalanced networks: A stochastic programming approach. *IISE Transactions* **49**(8) 781–799.

Fine, Charles H., Robert M. Freund. 1990. Optimal investment in product-flexible manufacturing capacity. *Management Science* **36**(4) 449–466.

Graves, Stephen C., Brian T. Tomlin. 2003. Process flexibility in supply chains. *Management Science* **49**(7) 907–919.

Hopp, Wallace J., Eylem Tekin, Mark P. Van Oyen. 2004. Benefits of skill chaining in serial production lines with cross-trained workers. *Management Science* **50**(1) 83–98.

Iravani, Seyed M., P. Van Oyen, Mark, Katharine T. Sims. 2005. Structural flexibility: A new perspective on the design of manufacturing and service operations. *Management Science* **51**(2) 151–166.

Iravani, Seyed M. R., Bora Kolfal, Mark P. Van Oyen. 2007. Call-center labor cross-training: Its a small world after all. *Management Science* **53**(7) 1102–1112.

Jordan, William C., Stephen C. Graves. 1995. Principles on the benefits of manufacturing process flexibility. *Management Science* **41**(4) 577–594.

Mak, Ho-Yin, Max Zuo-Jun Shen. 2009. Stochastic programming approach to process flexibility design. *Flexible Services and Manufacturing Journal* **21**(3-4) 75–91.

Shi, Cong, Yehua Wei, Yuan Zhong. 2019. Process flexibility for multiperiod production systems. *Operations Research* **67**(5) 1300–1320.

Simchi-Levi, David, He Wang, Yehua Wei. 2018. Increasing supply chain robustness through process flexibility and inventory. *Production and Operations Management* **27**(8) 1476–1491.

Simchi-Levi, David, Yehua Wei. 2012. Understanding the performance of the long chain and sparse designs in process flexibility. *Operations Research* **60**(5) 1125–1141.

Simchi-Levi, David, Yehua Wei. 2015. Worst-case analysis of process flexbility designs. *Operations Research* **63**(1) 166–185.

Tsitsiklis, John N., Kuang Xu. 2017. Flexible queueing architectures. *Operations Research* **65**(5) 1398–1413.

Van Mieghem, Jan A. 1998. Investment strategies for flexible resources. *Management Science* **44**(8) 1071–1078.

Wallace, Rodney B., Ward Whitt. 2005. A staffing algorithm for call centers with skill-based routing. *Manufacturing Service Operations Management* **7**(4) 276–294.

Wang, Shixin, Xuan Wang, Jiawei Zhang. 2019. A review of flexible processes and operations. Forthcoming at *Production and Operations Management*.

Wang, Shixin, Xuan Wang, Jiawei Zhang. 2020. Robust optimization approach to process flexibility designs with contribution margin differentials (unabridged version). Available at SSRN: `https://ssrn.com/abstract=3233721`.

Wang, Xuan, Jiawei Zhang. 2015. Process flexibility: A distribution-free bound on the performance of k-chain. *Operations Research* **63**(3) 555–571.

Electronic Companion to "Robust Optimization Approach to Process Flexibility Designs with Contribution Margin Differentials"

Shixin Wang, Xuan Wang, Jiawei Zhang

Appendix A: Proofs

Proof of Lemma 1. Since the dual problem (3) is a linear program, basic solutions are achieved at the corners of the polyhedron defined by $\{y_i + z_j = p_i \,|\, (a_j, b_i) \in \mathcal{A}\}$, $\{y_i = 0 \,|\, 1 \le i \le n\}$ and $\{z_j = 0 \,|\, 1 \le j \le m\}$. From the above $|\mathcal{A}| + m + n$ equations, we arbitrarily choose $m + n$ linearly independent binding constraints that identify a unique basic solution. In the chosen $m + n$ equations, denote C_1 as the set of equations that belong to $\{y_i = 0 \,|\, 1 \le i \le n\}$ or $\{z_j = 0 \,|\, 1 \le j \le m\}$, and denote C_2 as the set of equations that belong to $\{y_i + z_j = p_i \,|\, (a_j, b_i) \in \mathcal{A}\}$. Note that $\{y_i = 0 \,|\, 1 \le i \le n\} \cup \{z_j = 0 \,|\, 1 \le j \le m\}$ is not a feasible solution, and therefore we assume that $|C_2| > 0$. Let $G = (A, B, \mathcal{G})$ be the graph corresponding to the chosen $m + n$ equations, where the set of edges $\mathcal{G}$ is defined by C_2. In other words, $(a_j, b_i) \in \mathcal{G}$ if equation $y_i + z_j = p_i$ belongs to C_2. Suppose G consists of g components, each of which is a connected subgraph. For each subgraph $G_r = (A_r, B_r, \mathcal{G}_r)$ where $1 \le r \le g$, there are three cases to consider: (1) there exists some j such that $z_j = 0$; (2) there exists some i such that $y_i = 0$; and (3) $z_j \neq 0$ and $y_i \neq 0$ for all $a_j \in A_r, b_i \in B_r$.

Case 1: Suppose $z_{j^*} = 0$ for some $a_{j^*} \in A_r$. For any $a_{\hat{j}} \in A_r$, since G_r is connected, there exists a path $(a_{j^*}, b_{i_1}, a_{j_1}, b_{i_2}, \ldots, b_{i_k}, a_{\hat{j}})$ connecting a_{j^*} and $a_{\hat{j}}$. Since $y_i + z_j = p_i$ for any pair (i, j) on this path, $y_{i_1} = p_{i_1}, z_{j_1} = 0, \ldots, y_{i_{k-1}} = p_{i_{k-1}}, z_{j_{k-1}} = 0, y_{i_k} = p_{i_k}$ and $z_{\hat{j}} = 0$. Therefore, $z_j = 0$ for all $a_j \in A_r$. For any $b_{\hat{i}} \in B_r$, since G_r is connected, there exists a path $(a_{j^*}, b_{i_1}, a_{j_1}, \ldots, a_{j_k}, b_{\hat{i}})$ connecting a_{j^*} and $b_{\hat{i}}$. Since $y_i + z_j = p_i$ for any pair (i, j) on this path, $y_{i_1} = p_{i_1}$, $z_{j_1} = 0, \ldots, z_{j_k} = 0$, and $y_{\hat{i}} = p_{\hat{i}}$. Therefore, $y_i = p_i$ for any $b_i \in B_r$.

Case 2: Suppose $y_{i^*} = 0$ for some $b_{i^*} \in B_r$. For any $a_{\hat{j}} \in A_r$, since G_r is connected, there exists a path $(b_{i^*}, a_{j_1}, b_{i_1}, \ldots, b_{i_k}, a_{\hat{j}})$ connecting $a_{\hat{j}}$ and b_{i^*}. Since $y_i + z_j = p_i$ for any pair (i, j) on this path, $z_{j_1} = p_{i^*}, y_{i_1} = p_{i_1} - p_{i^*}, \ldots, z_{j_k} = p_{i^*}, y_{i_k} = p_{i_k} - p_{i^*}$, and $z_{\hat{j}} = p_{i^*}$. Therefore, $z_j = p_{i^*}$ for each $a_j \in A_r$. Similarly, there exists a path $(b_{i^*}, a_{j_1}, b_{i_1}, \ldots, a_{j_k}, b_{\hat{i}})$ connecting $b_{\hat{i}}$ and b_{i^*} for any $b_{\hat{i}} \in B_r$. Since $y_i + z_j = p_i$ for each pair (i, j) on this path, $z_{j_1} = p_{i^*}, y_{i_1} = p_{i_1} - p_{i^*}, \ldots, z_{j_k} = p_{i^*}$ and $y_{\hat{i}} = p_{\hat{i}} - p_{i^*}$. Therefore, $y_i = p_i - p_{i^*}$ for each $a_j \in A_r$. Notice $y_i + z_j = p_i$ for all $(a_j, b_i) \in G_r$, and hence b_{i^*} has the lowest margin and $p_i \ge p_{i^*}$ for all $b_i \in B_r$.

Case 3: Suppose $z_j \neq 0$ and $y_i \neq 0$ for all $a_j \in A_r, b_i \in B_r$. In this case, the chosen equations in G_r all come from subset C_2. Since there are $|A_r| + |B_r|$ variables, the number of arcs in G_r is also $|A_r| + |B_r|$. Notice that a connected graph with n vertices is a tree if it has $n - 1$ edges. It immediately follows that there must exist an even cycle in G_r since the number of the nodes is the same as that of the edges. But the arcs in an even cycle are linearly dependent, contradicting the fact that we have chosen $m+n$ linearly independent equations to identify a basic solution. $\square$

Proof of Theorem 1. To prove Theorem 1, we first show a result that focuses on the worst-case profit, and then we generalize to the broader class of robust measures in Γ.

PROPOSITION 3. *For any uncertainty set U that is part-wise independently symmetric around $\boldsymbol{\mu}$, the worst-case profit $R^g(\mathcal{A},U)$ is given by*

$$R^g(\mathcal{A},U)=\min_{K\in Q,\boldsymbol{\epsilon}\in E}\left\{\delta^K_{\boldsymbol{\mu}}(\mathcal{A})+\sum_{t=1}^{T}\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right)\right\}, \tag{EC.1}$$

where $\boldsymbol{\epsilon}_{s_t}:=\boldsymbol{d}_{s_t}-\boldsymbol{\mu}_{s_t}$ denotes the vector of demand residuals for each product category S_t.

Proof of Proposition 3. We first show that for any fixed $\boldsymbol{d}\in U$ and any given parameter set $K=\{k_{tr}\in\mathbb{Z}^+: t=1,...,T,\ r=0,...,t-1\}$, the worst-case profit $R^g(\mathcal{A},U)$ is bounded from above by the following:

$$R^g(\mathcal{A},U)\le\delta^K_{\boldsymbol{\mu}}(\mathcal{A})+\sum_{t=1}^{T}\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\min{}^i\boldsymbol{\epsilon}_{s_t}\right), \tag{EC.2}$$

where $\boldsymbol{\epsilon}_{s_t}:=\boldsymbol{d}_{s_t}-\boldsymbol{\mu}_{s_t}$ denotes the vector of demand residuals for each product category S_t. Then we shall show that there always exist some demand vector $\boldsymbol{d}^*$ and parameter set K^* such that the inequality in (EC.2) is tight, which in turn implies the representation (EC.1) of the worst-case profit.

To see (EC.2), notice that by the definition of $\delta^K_{\boldsymbol{\mu}}(\mathcal{A})$, there exist $\boldsymbol{y},\boldsymbol{z}$ that are feasible to problem (3) such that $\sum\limits_{i=1}^{n}\mu_i y_i+\sum\limits_{j=1}^{m}C_j z_j=\delta^K_{\boldsymbol{\mu}}(\mathcal{A})$, and $\sum_{b_i\in S_t}\mathbb{1}\{y_i=\beta_{tr}\}=k_{tr}$ for all $1\le t\le T, 0\le r\le t-1$. By the assumption on U, the uncertainty set for $\boldsymbol{\epsilon}_{s_t}$, denoted by $E_t:=\{\boldsymbol{\epsilon}_{s_t}|\boldsymbol{\epsilon}\in E\}$, is symmetric. Let σ^t be a permutation of the index set s_t such that $y_i=\beta_{tr}$ if and only if $\epsilon_{\sigma^t(i)}\in\{\min^i\boldsymbol{\epsilon}_{s_t}\,|\,H_{tr}-k_{tr}<i\le H_{tr}\}$. All such permutations σ^t for each s_t combined together form a permutation σ of $[n]$, and let $\boldsymbol{d}_\sigma=\boldsymbol{\mu}+\boldsymbol{\epsilon}_\sigma$. It then follows that

$$\sum_{i=1}^{n}d_{\sigma(i)}y_i+\sum_{j=1}^{m}C_j z_j=\delta^K_{\boldsymbol{\mu}}(\mathcal{A})+\sum_{i=1}^{n}\epsilon_{\sigma(i)}y_i=\delta^K_{\boldsymbol{\mu}}(\mathcal{A})+\sum_{t=1}^{T}\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\min{}^i\boldsymbol{\epsilon}_{s_t}\right). \tag{EC.3}$$

By definition, we have $g(\mathcal{A},\boldsymbol{d}_\sigma)\le\sum_{i=1}^{n}d_{\sigma(i)}y_i+\sum_{j=1}^{m}C_j z_j$. Since U is a part-wise independently symmetric perturbation uncertainty set, we have $\boldsymbol{\epsilon}_\sigma\in E$, which then implies that $R^g(\mathcal{A},U)\le g(\mathcal{A},d_\sigma)\le\delta^K_{\boldsymbol{\mu}}(\mathcal{A})+\sum_{t=1}^{T}\sum_{r=0}^{t-1}\beta_{tr}\left(\sum\limits_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\min{}^i\boldsymbol{\epsilon}_{s_t}\right)$.

Next, we prove there always exist some demand vector $\boldsymbol{d}^*$ and parameter set K^* such that the inequality in (EC.2) is tight. Let $\boldsymbol{d}^* = \arg\min_{\boldsymbol{d}\in U} g(\mathcal{A},\boldsymbol{d})$ and define $\boldsymbol{\tau} = \boldsymbol{d}^* - \boldsymbol{\mu}$. By definition, we have

$$g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\tau}) = \min_{\boldsymbol{y},\boldsymbol{z}} \sum_{i=1}^{n}(\mu_i+\tau_i)y_i + \sum_{j=1}^{m} C_j z_j \tag{EC.4}$$

$$\text{s.t. } y_i + z_j \geq p_i, \quad \forall (a_j, b_i) \in \mathcal{A},$$

$$y_i \geq 0, \forall 1 \leq i \leq n, \text{ and } z_j \geq 0, \forall 1 \leq j \leq m.$$

By Lemma 1, there exists an optimal solution $(\boldsymbol{y}^*, \boldsymbol{z}^*)$ to problem (EC.4) that satisfy $y_i^* \in \{\beta_{tr} \,|\, 0 \leq r \leq t\}$ for all $b_i \in S_t$ and $1 \leq t \leq T$. Let $k_{tr}^* = \sum_{b_i \in S_t} \mathbb{1}\{y_i^* = \beta_{tr}\}$ for all $0 \leq r \leq t-1$ and $1 \leq t \leq T$, and define $K^* = \{k_{tr}^* : t = 1, ..., T, \, r = 0, ..., t-1\}$. Let $H_{tr}^* := \sum_{j=0}^{r} k_{tj}^*$ for all $0 \leq r \leq t-1$ and $1 \leq t \leq T$. Then we must have $\sum_{i=1}^{n} y_i^* \tau_i \geq \sum_{t=1}^{T}\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}^*-k_{tr}^*+1}^{H_{tr}^*} \min{}^i \boldsymbol{\tau}_{s_t} \right)$ and $\sum_{j=1}^{m} C_j z_j^* + \sum_{i=1}^{n} \mu_i y_i^* \geq \delta_{\boldsymbol{\mu}}^{K^*}(\mathcal{A})$. It then follows that $R^g(\mathcal{A},U) = g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\tau}) \geq \delta_{\boldsymbol{\mu}}^{K^*}(\mathcal{A}) + \sum_{t=1}^{T}\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}^*-k_{tr}^*+1}^{H_{tr}^*} \min{}^i \boldsymbol{\tau}_{s_t} \right)$. On the other hand, it follows from (EC.2) that $R^g(\mathcal{A},U) = g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\tau}) \leq \delta_{\boldsymbol{\mu}}^{K^*}(\mathcal{A}) + \sum_{t=1}^{T}\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}^*-k_{tr}^*+1}^{H_{tr}^*} \min{}^i \boldsymbol{\tau}_{s_t} \right)$. Therefore, we have

$$R^g(\mathcal{A},U) = \delta_{\boldsymbol{\mu}}^{K^*}(\mathcal{A}) + \sum_{t=1}^{T}\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}^*-k_{tr}^*+1}^{H_{tr}^*} \min{}^i \boldsymbol{\tau}_{s_t} \right) \tag{EC.5}$$

Now we are ready to complete the proof of (EC.1). By (EC.2) and (EC.5), we get

$$\begin{aligned}
R^g(\mathcal{A},U) &= \min_{K\in Q} \left\{ \delta_{\boldsymbol{\mu}}^{K}(\mathcal{A}) + \sum_{t=1}^{T} \min_{\boldsymbol{\epsilon}_{s_t}\in E_t} \left(\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}} \min{}^i \boldsymbol{\epsilon}_{s_t} \right)\right)\right\} \\
&= \min_{K\in Q} \left\{ \delta_{\boldsymbol{\mu}}^{K}(\mathcal{A}) + \sum_{t=1}^{T} \min_{\boldsymbol{\epsilon}_{s_t}\in E_t} \left(\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}} \epsilon_{s_t,i} \right)\right)\right\} \\
&= \min_{K\in Q, \boldsymbol{\epsilon}_{s_t}\in E_t, \forall 1\leq t\leq T} \left\{ \delta_{\boldsymbol{\mu}}^{K}(\mathcal{A}) + \sum_{t=1}^{T} \left(\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}} \epsilon_{s_t,i} \right)\right)\right\} \\
&= \min_{K\in Q, \boldsymbol{\epsilon}\in E} \left\{ \delta_{\boldsymbol{\mu}}^{K}(\mathcal{A}) + \sum_{t=1}^{T} \left(\sum_{r=0}^{t-1} \beta_{tr} \left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}} \epsilon_{s_t,i} \right)\right)\right\},
\end{aligned}$$

where the second equality holds because E is part-wise independently symmetric. □

We now complete the proof of Theorem 1. For each $1 \leq t \leq T$, let $E_t^{w_t} := \{\boldsymbol{\epsilon}_{s_t} \in E_t \,|\, \sum \boldsymbol{\epsilon}_{s_t} = w_t\}$ for any given $w_t \in \mathbb{R}_+$. Note that both $E_t^{w_t}$ and $(E_1^{w_1} \times E_2^{w_2} \times ... \times E_T^{w_T}) \bigcap E$ are part-wise independently symmetric. For notation brevity, denote $\prod_{t=1}^{T} E_t^{w_t} = E_1^{w_1} \times E_2^{w_2} \times ... \times E_T^{w_T}$. Now we have

$$R^f(\mathcal{A},U) = \min_{\boldsymbol{d}\in U} \left\{ h\left(g(\mathcal{A},\boldsymbol{d}), \sum_{i\in s_1} d_i, ..., \sum_{i\in s_T} d_i \right)\right\} = \min_{\boldsymbol{\epsilon}\in E} \left\{ h\left(g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\epsilon}), \sum_{i\in s_1}(\mu_i+\epsilon_i), ..., \sum_{i\in s_T}(\mu_i+\epsilon_i) \right)\right\}$$

$$= \min_{w_t, \forall 1\le t\le T}\left\{\min_{\boldsymbol{\epsilon}\in\prod_{t=1}^T E_t^{w_t}\bigcap E}\left\{h\left(g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\epsilon}),\sum_{i\in s_1}(\mu_i+\epsilon_i),...,\sum_{i\in s_T}(\mu_i+\epsilon_i)\right)\right\}\right\}$$

$$= \min_{w_t, \forall 1\le t\le T}\left\{\min_{\boldsymbol{\epsilon}\in\prod_{t=1}^T E_t^{w_t}\bigcap E}\left\{h\left(g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\epsilon}),\sum_{i\in s_1}\mu_i+w_1,...,\sum_{i\in s_T}\mu_i+w_T\right)\right\}\right\}.$$

Since h is strictly increasing in $g(\mathcal{A},\boldsymbol{d})$ for fixed $w_t, \forall t=1,...,T$, we have $R^f(\mathcal{A},U) = \min_{w_t,\forall 1\le t\le T}\left\{h\left(\min_{\boldsymbol{\epsilon}\in\prod_{t=1}^T E_t^{w_t}\bigcap E} g(\mathcal{A},\boldsymbol{\mu}+\boldsymbol{\epsilon}),\sum_{i\in s_1}\mu_i+w_1,...,\sum_{i\in s_T}\mu_i+w_T\right)\right\}$. By Proposition 3,

$$R^f(\mathcal{A},U) = \min_{w_t,\forall 1\le t\le T}\left\{h\left(\min_{K\in Q,\boldsymbol{\epsilon}\in\prod_{t=1}^T E_t^{w_t}\bigcap E}\delta_{\boldsymbol{\mu}}^K(\mathcal{A})+\sum_{t=1}^T\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right),\sum_{i\in s_1}\mu_i+w_1,...,\sum_{i\in s_T}\mu_i+w_T\right)\right\}$$

$$= \min_{w_t,\forall 1\le t\le T, K\in Q,\boldsymbol{\epsilon}\in\prod_{t=1}^T E_t^{w_t}\bigcap E}\left\{h\left(\delta_{\boldsymbol{\mu}}^K(\mathcal{A})+\sum_{t=1}^T\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right),\sum_{i\in s_1}\mu_i+w_1,...,\sum_{i\in s_T}\mu_i+w_T\right)\right\}$$

$$= \min_{K\in Q,\boldsymbol{\epsilon}\in E} h\left(\delta_{\boldsymbol{\mu}}^K(\mathcal{A})+\sum_{t=1}^T\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right),\sum_{i\in s_1}(\mu_i+\epsilon_i),\sum_{i\in s_2}(\mu_i+\epsilon_i),...,\sum_{i\in s_T}(\mu_i+\epsilon_i)\right),$$

which completes the proof of Theorem 1. □

Proof of Theorem 2. (a). Let f be the deterministic measure function of R and $h(\cdot)$ be the function such that $h(x,y_1,...,y_T)$ is strictly increasing in x for fixed $\{y_1,...,y_T\}$. For each $l=1,2$, let $f(\mathcal{A}_l,\boldsymbol{d}) = h(g(\mathcal{A}_l,\boldsymbol{d}),\sum\boldsymbol{d}_{s_1},...,\sum\boldsymbol{d}_{s_T})$. Then by Theorem 1, we have

$$R^f(\mathcal{A}_l,U) = \min_{K\in Q,\boldsymbol{\epsilon}\in E} h\left(\delta_{\boldsymbol{\mu}}^K(\mathcal{A}_l)+\sum_{t=1}^T\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right),\sum_{i\in s_1}(\mu_i+\epsilon_i),...,\sum_{i\in s_T}(\mu_i+\epsilon_i)\right).$$

If $\delta_{\boldsymbol{\mu}}^K(\mathcal{A}_1)\ge\delta_{\boldsymbol{\mu}}^K(\mathcal{A}_2)$ for all $K\in Q$, it then follows from the above equation that $R(\mathcal{A}_1,U)\ge R(\mathcal{A}_2,U)$.

(b). Let $W^{k_1,...,k_T}\in Q$ denote the parameter set whose only non-negative entries are at $(t,0)$ with value k_t for each $1\le t\le T$. That is, $W_{t0}^{k_1,...,k_T}=k_t$ and $W_{tr}^{k_1,...,k_T}=0$ for all $r>0$. Suppose there exist some $\{k_1,...,k_T\}$ such that $\delta_{\boldsymbol{\mu}}^{W^{k_1,...,k_T}}(\mathcal{A}_1)<\delta_{\boldsymbol{\mu}}^{W^{k_1,...,k_T}}(\mathcal{A}_2)$. We next construct a part-wise independently symmetric perturbation uncertainty set U^* such that $R(\mathcal{A}_1,U^*)<R(\mathcal{A}_2,U^*)$, which contradicts with the assumption that $R(\mathcal{A}_1,U)\ge R(\mathcal{A}_2,U)$ for any part-wise independently symmetric perturbation uncertainty set U and hence would complete the proof. Let $M=\frac{\max_{K^1,K^2\in Q}|\delta_{\boldsymbol{\mu}}^{K^1}(\mathcal{A}_2)-\delta_{\boldsymbol{\mu}}^{K^2}(\mathcal{A}_2)|}{\min_{1\le t\le T,0\le r<t}\beta_{tr}}$. Let $\boldsymbol{\epsilon}_{s_t}^*$ be a $|S_t|$-dimensional vector such that there are exactly k_t entries with value equal to $-M$, and the other $|S_t|-k_t$ entries are equal to M. Define $E^*:=\prod_{t=1}^T(\Omega(\boldsymbol{\epsilon}_{s_t}^*))$, where $\Omega(\boldsymbol{\epsilon}_{s_t}^*)$ is the set of all vectors that are permutations of $\boldsymbol{\epsilon}_{s_t}^*$ restricted to product subset S_t. Let $U^*=\{\boldsymbol{\mu}+\boldsymbol{\epsilon}|\boldsymbol{\epsilon}\in E^*\}$. By our choice of U^* and the definition of M, we have

$$R(\mathcal{A}_2,U^*) = \min_{K\in Q,\boldsymbol{\epsilon}\in E^*} h\left(\delta_{\boldsymbol{\mu}}^K(\mathcal{A}_2)+\sum_{t=1}^T\left(\sum_{r=0}^{t-1}\beta_{tr}\left(\sum_{i=H_{tr}-k_{tr}+1}^{H_{tr}}\epsilon_{s_t,i}\right)\right),\sum_{i\in s_1}(\mu_i+\epsilon_i),...,\sum_{i\in s_T}(\mu_i+\epsilon_i)\right)$$

$$
\begin{aligned}
&= h\left(\delta_{\boldsymbol{\mu}}^{W^{k_1,\dots,k_T}}(\mathcal{A}_2) - \sum_{t=1}^{T} P_t k_t M, \sum_{i\in s_1}\mu_i + (|S_1| - 2k_1)M, \dots, \sum_{i\in s_T}\mu_i + (|S_T| - 2k_T)M\right)\\
&> h\left(\delta_{\boldsymbol{\mu}}^{W^{k_1,\dots,k_T}}(\mathcal{A}_1) - \sum_{t=1}^{T} P_t k_t M, \sum_{i\in s_1}\mu_i + (|S_1| - 2k_1)M, \dots, \sum_{i\in s_T}\mu_i + (|S_T| - 2k_T)M\right)\\
&\geq R(\mathcal{A}_1, U^*),
\end{aligned}
$$

which completes the proof of part (b). □

Proof of Lemma 2. We analyze problem (5) for the alternate long chain design in three cases. Recall that $k_{10} = \{b_i \in S_L \,|\, y_i = P_L\}$, $k_{20} = \{b_i \in S_H \,|\, y_i = P_H\}$, and $k_{21} = \{b_i \in S_H \,|\, y_i = P_H - P_L\}$.

Case 1: $k_{20} = 0$**.** In this case, the y_i value associated with each high profit product $b_i \in S_H$ is equal to either 0 or $P_H - P_L$. Consider an optimal solution $\boldsymbol{z}^*$ to problem (5). For each $b_i \in S_H$ with $y_i = 0$, we have $z_j^* = P_H$ for its two neighboring plant nodes $j \in N(b_i, \mathcal{AL})$. By the specific structure of the alternate long chain that each plant is connected to exactly one high profit product and one low profit product, we have $|\{a_j \in A \,|\, z_j^* = P_H\}| = 2k_{22}$. For the rest of high profit products with $y_i = P_H - P_L$, the smallest value of z_j for their neighboring plant nodes that guarantees feasibility of problem (5) is $z_j^* = P_L$. Therefore, the optimal objective value of (5) is equal to $2k_{22}P_H + 2k_{21}P_L$.

Case 2: $k_{20} = n/2$**.** In this case, all the $n/2$ high profit products have their y_i value equal to P_H. Consider an optimal solution $\boldsymbol{z}^*$ to problem (5). The feasibility requirement for a strictly positive z_j^* value solely comes from the low profit products whose associated y_i is equal to zero, in which case its neighboring plant nodes have $z_j^* = P_L$. In the alternate long chain design, $N(b_i, \mathcal{AL}) \cap N(b_j, \mathcal{AL}) = \emptyset$ for any $b_i, b_j \in S_L$ such that $i \neq j$. Therefore, the optimal objective value of (5) is equal to $2k_{11}P_L$.

Case 3: $k_{20} = 1, \dots, n/2 - 1$**.** In this case, we can spell out the expression in (6) as follows:

$$
\delta^K(\mathcal{AL}) = \begin{cases} 2k_{22}P_H + 2k_{21}P_L, & \text{if } k_{10} \geq k_{20} + 1 \\ 2k_{22}P_H + (2k_{21}+1)P_L, & \text{if } k_{10} = k_{20}, \\ 2k_{22}P_H + (2k_{11} - 2k_{22})P_L, & \text{if } k_{10} \leq k_{20} - 1, \end{cases} \tag{EC.6}
$$

We next provide some high-level idea about the analysis of problem (5), and then break down the analysis into three subcases and show that the optimal objective value of (5) is the same as (EC.6).

Notice that with parameter set K given, minimizing objective function (5) reduces to finding an assignment of the y_i values to product nodes such that the feasibility requirement on z_j (specified by $y_i + z_j \geq p_i$ for each arc $(a_j, b_i) \in \mathcal{AL}$) can be achieved by the smallest value possible. It is easy to see that there exists an optimal solution to (5) such that $z_j^* \in \{0, P_L, P_H\}$ for any $1 \le j \le n$. Second, for each one of the k_{22} high profit products with $y_i = 0$, all their distinct $2k_{22}$ neighboring plant nodes must have $z_j^* = P_H$. Finally, for each plant j, the only case that z_j^* can take value zero happens when $y_i = p_i$ for each $b_i \in N(a_j, \mathcal{AL})$. Therefore, in order to minimize $\sum_j z_j$, it is optimal to assign the k_{10} low profit products with $y_i = P_L$ and the k_{20} high profit products with $y_i = P_H$ in a consecutive manner to the greatest extent possible so as to maximize the number of plant nodes with $z_j^* = 0$. We consider three subcases as follows.

(a) If $k_{10} \ge k_{20} + 1$, then the maximum number of consecutive products with $y_i = p_i$ is $2k_{20} + 1$, and the maximum number of plant nodes with $z_j^* = 0$ is $2k_{20}$. We next show there does exist a feasible solution such that the number of plant nodes with $z_j^* = 0$ is $2k_{20}$, which implies that the number of plant nodes with $z_j^* = P_L$ is $n - 2k_{22} - 2k_{20} = 2k_{21}$ and therefore $\delta^K(\mathcal{AL}) = 2k_{22}P_H + 2k_{21}P_L$. Consider an alternate long chain $\mathcal{AL}$ with $p_1 = p_3 = \cdots = p_{n-1} = P_H$ and $p_2 = p_4 = \cdots = p_n = P_L$. Let $y_i = 0$ for $i \in \{1, 3, \ldots, 2k_{22} - 1\} \cup \{2, 4, \ldots, 2k_{11}\}$; let $y_i = P_L$ for $i \in \{2k_{11} + 2, \ldots, n\}$; let $y_i = P_H - P_L$ for $i \in \{2k_{22} + 1, \ldots, n - 2k_{20} - 1\}$; and let $y_i = P_H$ for $i \in \{n - 2k_{20} + 1, \ldots, n\}$. Then we have $z_j = P_H$ for all $j \in \{1, \ldots, 2k_{22}\}$, $z_j = P_L$ for all $j \in \{2k_{22} + 1, \ldots, 2k_{22} + 2k_{21}\}$ and $z_j = 0$ for all $j \in \{2k_{22} + 2k_{21} + 1, \ldots, n\}$, and the number of plant nodes with $z_j = 0$ is $n - 2k_{22} - 2k_{21} = 2k_{20}$.

(b) If $k_{10} = k_{20}$, then the maximum number of consecutive products with $y_i = p_i$ is $2k_{10}$, and hence the maximum number of plant nodes with $z_j^* = 0$ is $2k_{20} - 1$. Similar to the analysis in (a), we next show there does exist a feasible solution such that the number of plant nodes with $z_j^* = 0$ is $2k_{20} - 1$, which then implies that the number of plant nodes with $z_j^* = P_L$ is $n - 2k_{22} - 2k_{20} + 1 = 2k_{21} + 1$ and hence $\delta^K(\mathcal{AL}) = 2k_{22}P_H + (2k_{21} + 1)P_L$. Consider the same design $\mathcal{AL}$ as in (a). Let $y_i = 0$ for $i \in \{1, 3, \ldots, 2k_{22} - 1\} \cup \{2, 4, \ldots, 2k_{11}\}$; let $y_i = P_L$ for $i \in \{n - 2k_{10} + 2, \ldots, n\}$; let $y_i = P_H - P_L$ for $i \in \{2k_{22} + 1, \ldots, n - 2k_{20} - 1\}$; and let $y_i = P_H$ for $i \in \{n - 2k_{20} + 1, \ldots, n\}$. Then we have $z_j = P_H$ for all $j \in \{1, \ldots, 2k_{22}\}$, $z_j = P_L$ for all $j \in \{2k_{22} + 1, \ldots, 2k_{22} + 2k_{21} + 1\}$ and $z_j = 0$ for all $j \in \{2k_{22} + 2k_{21} + 2, \ldots, n\}$, in which case the number of plant nodes with $z_j = 0$ is $n - 2k_{22} - 2k_{21} - 1 = 2k_{20} - 1$.

(c) If $k_{10} \le k_{20} - 1$, then the maximum number of consecutive products with $y_i = p_i$ is $2k_{10} + 1$. Consider the same design $\mathcal{AL}$ as in (a) and (b). Following a similar analysis, it is straightforward to show that there exist a feasible solution with $z_j = P_H$ for all $j \in \{1, \ldots, 2k_{22}\}$, $z_j = P_L$ for all $j \in \{2k_{22} + 1, \ldots, 2k_{11}\}$ and $z_j = 0$ for all $j \in \{2k_{11} + 2, \ldots, n\}$, which has the maximum possible number $2k_{10}$ of plants with

$z_j^* = 0$. It then follows that the number of plants with $z_j^* = P_L$ is $n - 2k_{22} - 2k_{10} = 2k_{11} - 2k_{22}$, which implies $\delta^K(\mathcal{AL}) = 2k_{22}P_H + (2k_{11} - 2k_{22})P_L$. $\square$

Proof of Proposition 1. Consider a long chain with $n/2$ low profit products and $n/2$ high profit products, where n is an even integer. We call a consecutive sequence of k product nodes a *sequence of length k*. We first provide some structural properties of sequences with an odd (cf. Lemma 3(a)) and even (cf. Lemma 3(b)) number of product nodes. We refer the readers to Wang et al. (2020) Appendix A for the proof of Lemma 3.

LEMMA 3. *(a) For any $1 \le k \le n/2 - 1$, there exists a sequence of length $2k+1$ with k high profit products and $k+1$ low profit products. (b) For any $1 \le k \le n/2$, there exists a sequence of length $2k$ with k high profit products and k low profit products.*

We next prove Proposition 1, which together with Theorem 2(a) implies Theorem 3. Consider a long chain $\mathcal{L}_{HL}$ with $n/2$ high profit products and $n/2$ low profit products. Consider a fixed $K = \{k_{10}, k_{20}, k_{21}\}$. Recall that $k_{11} = n/2 - k_{10}$ is equal to $|\{b_i \in S_L \mid y_i = 0\}|$, and $k_{22} = n/2 - k_{20} - k_{21}$ is equal to $|\{b_i \in S_H \mid y_i = 0\}|$.

Case 1: $k_{20} = 0$. In this case, since every product is connected to exactly 2 plants, for any optimal solution $\boldsymbol{z}^*$ to (5) of design $\mathcal{L}_{HL}$, the number of z_j^* such that $z_j^* = P_H$ is at most $2k_{22}$. The value of the other z_j^*'s except the above (at most) $2k_{22}$ nodes is either P_L or 0 since these plants are not connected to high profit products with $y_i = 0$. By Lemma 2, it then follows that $\delta^K(\mathcal{L}_{HL}) \le 2k_{22}P_H + (n - 2k_{22})P_L = 2k_{22}P_H + 2k_{21}P_L = \delta^K(\mathcal{AL})$.

Case 2: $k_{20} = n/2$. In this case, the y_i values of all high profit products are equal to P_H and therefore the feasibility requirement for a strictly positive z_j value solely comes from the low profit products whose associated y_i value is zero, in which case $z_j^* = P_L$. Since every product node is connected to exactly 2 plant nodes and $|\{b_i \in S_L \mid y_i = 0\}| = k_{11}$, we have $\delta^K(\mathcal{L}_{HL}) \le 2k_{11}P_L = \delta^K(\mathcal{AL})$.

Case 3: $k_{20} = 1, ..., n/2 - 1$. In this case, the analysis depends on the relationship between k_{10} and k_{20}.

(a) If $k_{10} \ge k_{20} + 1$: Since $k_{20} \le n/2 - 1$, by Lemma 3(a), there exists a sequence $\mathcal{S}_a$ of length $2k_{20} + 1$ with k_{20} high profit products and $k_{20} + 1$ low profit products. Let $y_i = P_H$ if product i belongs to $\mathcal{S}_a$ and $p_i = P_H$. Similarly, let $y_i = P_L$ if product i belongs to $\mathcal{S}_a$ and $p_i = P_L$. Therefore, for an optimal solution $\boldsymbol{z}^*$ to (5) of design $\mathcal{L}_{HL}$, the $2k_{20}$ plants that are solely connected to the above $2k_{20} + 1$ products in $\mathcal{S}_a$ all have $z_j^* = 0$. This implies that among all the n plants, the number of plants whose z_j^* value is equal to zero is at least $2k_{20}$. Therefore, the number of plants whose z_j^* value is equal to either P_L or P_H is at most $n - 2k_{20}$. Since each product has exactly two neighboring plants, the number of plants with $z_j^* = P_H$ is at most $2k_{22}$. Therefore, $\delta^K(\mathcal{L}_{HL}) \le 2k_{22}P_H + (n - 2k_{20} - 2k_{22})P_L = 2k_{22}P_H + 2k_{21}P_L = \delta^K(\mathcal{AL})$.

(b) If $k_{10} = k_{20}$: Since $k_{20} \leq n/2$, by Lemma 3(b), there exists a sequence $\mathcal{S}_b$ of length $2k_{20}$ with k_{20} high profit products and k_{20} low profit products. Let $y_i = P_H$ if product i belongs to $\mathcal{S}_b$ and $p_i = P_H$. Let $y_i = P_L$ if product i belongs to $\mathcal{S}_b$ and $p_i = P_L$. Therefore, for an optimal solution $\boldsymbol{z}^*$ to (5) of design $\mathcal{L}_{HL}$, the $2k_{20} - 1$ plants that are solely connected to the above $2k_{20}$ products in $\mathcal{S}_b$ all have $z_j^* = 0$. This implies that among all the n plants, the number of plants whose z_j^* value is equal to zero is at least $2k_{20} - 1$. Therefore, the number of plants whose z_j^* value is equal to either P_L or P_H is at most $n - 2k_{20} + 1$. Since each product has exactly two neighboring plants, the number of plants with $z_j^* = P_H$ is at most $2k_{22}$. It then follows that $\delta^K(\mathcal{L}_{HL}) \leq 2k_{22}P_H + (n - 2k_{20} - 2k_{22} + 1)P_L = 2k_{22}P_H + (2k_{21} + 1)P_L = \delta^K(\mathcal{AL})$.

(c) If $k_{10} \leq k_{20} - 1$: In this case, we have $k_{10} \leq n/2 - 1$. By Lemma 3(a), there exists a sequence $\mathcal{S}_c$ of length $2k_{10} + 1$ with k_{10} high profit products and $k_{10} + 1$ low profit products. Let $y_i = P_H$ if product i belongs to $\mathcal{S}_c$ and $p_i = P_H$. Let $y_i = P_L$ if product i belongs to $\mathcal{S}_c$ and $p_i = P_L$. Therefore, for an optimal solution $\boldsymbol{z}^*$ to (5) of design $\mathcal{L}_{HL}$, the $2k_{10}$ plants that are solely connected to the above $2k_{10} + 1$ products in $\mathcal{S}_c$ all have $z_j^* = 0$. This implies that among all the n plants, the number of plants whose z_j^* value is equal to zero is at least $2k_{10}$. Therefore, the number of plants whose z_j^* value is equal to either P_L or P_H is at most $n - 2k_{10}$. Since each product has exactly two neighboring plants, the number of plants with $z_j^* = P_H$ is at most $2k_{22}$. Therefore, $\delta^K(\mathcal{L}_{HL}) \leq= 2k_{22}P_H + (n - 2k_{10} - 2k_{22})P_L = 2k_{22}P_H + (2k_{11} - 2k_{22})P_L = \delta^K(\mathcal{AL})$.

Combining all the above three cases together with Theorem 2(a) completes the proof. □

Proof of Proposition 2. We show there exist parameter sets $K = \{k_{10}, k_{20}, k_{21}\}$ and $K' = \{k'_{10}, k'_{20}, k'_{21}\}$ with $k_{21} = k'_{21} = 0$ such that $\delta^K(\mathcal{AL}) > \delta^K(\mathcal{DL})$ and $\delta^{K'}(\mathcal{AL}) < \delta^{K'}(\mathcal{DL})$, which implies Proposition 2 by Theorem 2(b). The disjoint long chain $\mathcal{DL}$ consists of two disconnected components, denoted by $\mathcal{DL}_H$ and $\mathcal{DL}_L$, where each component is a long chain with size $n/2$ and the products in component $\mathcal{DL}_H$ (resp. $\mathcal{DL}_L$) have identical margins P_H (resp. P_L). By the definition of DMGI, we have $\delta^K(\mathcal{DL}) = \delta^K(\mathcal{DL}_H) + \delta^K(\mathcal{DL}_L)$, where $\delta^K(\mathcal{DL}_H)$ and $\delta^K(\mathcal{DL}_L)$ are the optimal objective values to problem (5) for designs $\mathcal{DL}_H$ and $\mathcal{DL}_L$. Consider the following two parameter sets K and K':

- $K = \{k_{10}, k_{20}, k_{21}\}$ with $k_{10} = \frac{n}{2}, k_{20} = k_{21} = 0$. In this case, we have $k_{22} = \frac{n}{2}$ and it follows from Lemma 2 that $\delta^K(\mathcal{AL}) = nP_H$. For the disjoint long chain, it is easy to see that $\delta^K(\mathcal{DL}_H) = \frac{n}{2}P_H$ and $\delta^K(\mathcal{DL}_L) = 0$, and hence $\delta^K(\mathcal{DL}) = \frac{n}{2}P_H$. Therefore, we have $\delta^K(\mathcal{AL}) > \delta^K(\mathcal{DL})$.

- $K' = \{k'_{10}, k'_{20}, k'_{21}\}$ with $k'_{10} = \frac{n}{2} - 1, k'_{20} = \frac{n}{2} - 1, k'_{21} = 0$. In this case, we have $k'_{22} = 1$ and by Lemma 2, we have $\delta^{K'}(\mathcal{AL}) = 2P_H + \left(n - 2 - \min\left(\frac{n}{2} - 1, \frac{n}{2}\right) - \min\left(\frac{n}{2} - 1, \frac{n}{2} - 2\right)\right)P_L = 2P_H + P_L$. Now we consider the disjoint long chain. For component $\mathcal{DL}_L$, we have $k'_{11} = \frac{n}{2} - k'_{10} = 1$ and hence $\delta^{K'}(\mathcal{DL}_L) = (k'_{11} + 1)P_L = 2P_L$. For component $\mathcal{DL}_H$, we have $k'_{22} = \frac{n}{2} - k'_{20} - k'_{21} = 1$ and $\delta^{K'}(\mathcal{DL}_H) = (k'_{22} + 1)P_H = 2P_H$. It then follows that $\delta^{K'}(\mathcal{DL}) = 2P_H + 2P_L$, and therefore we have $\delta^{K'}(\mathcal{AL}) < \delta^{K'}(\mathcal{DL})$. $\square$