Product Flexibility Strategy Under Supply and Demand Risk

Yimin Wang, Scott Webster
W. P. Carey School of Business, PO Box 874706, Tempe, AZ 85287-4706, yimin_wang@asu.edu, scott.webster@asu.edu

Problem definition: With heightened global uncertainty, supply chain managers are under increasing pressure to craft strategies that accommodate both supply and demand risks. While product flexibility is a well-understood strategy to accommodate risk, there is no clear guidance on the optimal flexibility configuration of a supply network that comprises both unreliable primary suppliers and reliable backup suppliers. **Academic/Practical relevance:** Existing literature examines the value of flexibility with primary and backup suppliers independently. For a risk-neutral firm, research shows that (a) incorporating flexibility in a primary supplier by replacing two dedicated ones (in absence of backup supply) is always beneficial, and (b) adding flexibility to a reliable backup supplier (in absence of product flexibility in primary suppliers) is always valuable. It is unclear, however, how flexibility should be incorporated into a supply network with both unreliable primary suppliers and reliable backup suppliers. This research studies whether flexibility should be incorporated in a primary supplier, a backup supplier, or both. **Methodology:** We develop a normative model to analyze when flexibility benefits and when it hurts. **Results:** Compared with a base case of no flexibility, we prove that incorporating flexibility in either primary or backup suppliers is always beneficial. However, incorporating flexibility in both primary and backup suppliers can be counterproductive because the supply chain performance can decline with saturated flexibility, even if flexibility is costless. A key reason is that the risk-aggregation effect of consolidating flexibility in an unreliable supplier becomes more salient when flexibility is already embedded in a backup supplier. **Managerial implications:** This research refines the existing understanding of flexibility by illustrating that flexibility is not always beneficial. When there is a choice, a firm should prioritize incorporating flexibility in a reliable backup supplier.

Key words: flexibility, risk, supply network

1. Introduction

Developing a resilient and flexible supply chain is an imperative task in today's business environment. As both supply and demand risks increase, firms start to re-examine their supply chain configurations: 64% of respondents to PwC's recent Global Supply Chain Survey said they plan to implement greater product flexibility (i.e., the ability for a supplier to produce multiple products) to better respond to demand uncertainty, while at the same time they also recognized that supply chains are increasingly vulnerable to disruptions caused by natural disasters, political unrest, and port strikes (Sun and Goldbach 2011). Since improved flexibility may not offer much value if the supply chain is vulnerable to disruptions, the issue of supply chain resilience must be addressed concurrently with flexibility.

A common strategy firms employ to improve supply chain resilience is to contract with a backup supplier as well as with a primary vendor. This strategy allows a firm to have an alternative supplier that could quickly fill the void if the key vendor fails to deliver. The use of backup suppliers occurs in many industries, including the automotive (Sheffi 2005), pharmaceuticals (Chopra and Sodhi 2004), global manufacturing (Huchzermeier and Cohen 2006), and electronics industries (Tomlin 2006). Toyota, for example, uses backup suppliers to reduce its exposure to disruptions (Kim 2011). Likewise, GE Aviation contracts with two certified suppliers, but specifies one as a primary vendor and the other as a backup that fills orders when the primary vendor cannot meet GE's demand (Zeng and Xia 2015). Kouvelis and Li (2008) note that using backup suppliers is particularly suited for just-in-time production environments, where the most frequent cause of disruptions is unreliable supply. Sting and Huchzermeier (2010) and Spinler and Huchzermeier (2006) show that it can be Pareto improving for both parties. In sum, supply networks that consists of both primary and backup suppliers are widely implemented in practice.

Using backup suppliers improves a supply chain's resilience to supply-side disruptions, but this strategy does not address demand-side risks. Flexibility, on the other hand, can effectively address the demand-side risk. From an operational point of view, flexibility can be added to a supply network by either replacing multiple product-dedicated suppliers with a flexible supplier that is capable of supplying multiple products or by converting an existing supplier into a flexible one through supplier development efforts. An important operational benefit of flexibility is that fewer suppliers are required to produce multiple products; as such, one flexible supplier can displace multiple dedicated suppliers. In this study, we focus on the scenario where 'embed' or 'incorporate' flexibility means that a *single flexible supplier displaces multiple dedicated suppliers*. This approach is in line with the large stream of literature that allows flexibility to displace dedicated suppliers either explicitly (Baker et al. 1986, Tomlin and Wang 2005, Thonemann and Brandeau 2000, Chod et al. 2021) or implicitly (Eppen et al. 1989, Fine and Freund 1990, Van Mieghem 1998, 2004). Note that an alternative scenario is that flexibility does not displace any existing suppliers (Jordan and Graves 1995, Bassamboo et al. 2010). This line of research assumes flexibility is always valuable and reliable, thus it focuses on the marginal value of flexibility. In contrast, we do not assume flexibility is always reliable, hence flexibility may not be always beneficial.

This study investigates whether flexibility can best address supply and demand risk when it is embedded in a primary supplier, in a backup supplier, or simultaneously in both suppliers. Suppose a firm's primary supplier is unreliable but cheap, while the firm's backup supplier is reliable but more expensive than the primarily supplier. It is not immediately clear how flexibility should be embedded in such a supply network. The firm could incorporate flexibility through a flexible primary supplier, which would offer a less expensive cost for all products, but it would also suffer from risk-aggregation disadvantage, which would eliminate flexibility advantage if it fails. In contrast, the firm could incorporate flexibility by contracting with a flexible backup supplier to offer reliable supplies for multiple products, although this option would incur a higher unit cost for all products. Intuitively, either of these two alternative strategies could dominate, depending on the magnitude of supply versus demand risks.

If the above intuition is correct, then a natural question is whether a firm can take advantage of the best of both worlds by embedding flexibility in both primary and backup suppliers, constructing a fully flexible supply network. If incorporating flexibility is free, then the answer – according to extant studies – appears to be affirmative: embedding flexibility cannot hurt supply chain performance. Existing literature supports the idea that incorporating flexibility is always beneficial. For a risk-neutral firm, research shows that (a) incorporating flexibility in an unreliable primary supplier (in absence of backup supply) is beneficial, and (b) incorporating flexibility in a reliable backup supplier (in absence of flexibility in primary suppliers) is also valuable. These results hint that pursuing a fully flexible supply network can never hurt.

Using a basic two-product supply network, we find that, surprisingly, the above reasoning is not true in general. There exist a wide range of supply and demand risk profiles in which achieving a fully flexible supply network is not beneficial. This finding is perplexing and counter intuitive since existing literature shows that flexibility can not hurt supply chain performance – although the marginal benefit declines with increased flexibility (Bassamboo et al. 2010, 2012). The reason why flexibility may hurt supply chain performance can be understood by examining a supply network where there are no reliable backup suppliers. In this case, embedding flexibility in a supplier always improves supply chain performance, regardless of how unreliable the supplier is. This result is proved in Tomlin and Wang (2005), but the intuition is not straightforward because, one expects that if a supplier is highly unreliable, then it is unwise to rely on this supplier for different products since the supply for both products will be lost in case of a disruption. Instead, a dedicated supplier for each product line seems more attractive since this approach significantly reduces the chance that the supply to both products will be lost. Hence, embedding flexibility into an unreliable supplier has a risk-aggregation disadvantage (both products are unavailable in the case of a disruption), but it offers a demand-pooling advantage in the absence of a disruption. Why, then, does the demand-pooling benefit always dominate the risk-aggregation disadvantage (as shown in Tomlin and Wang 2005)?

The answer to the above question is that a flexibility-endowed supplier can replicate any strategy adopted by product-dedicated suppliers. To be concrete, let *total failure* denote the instances where both products fail; and *partial failure* denote the instances where only one product fails. As an example, a partial failure occurs when the supplier to product 1 is down but the supplier to product 2 is up. If each supplier has a failure probability ρ, then this partial failure occurs with a probability $\rho(1-\rho)$. On the surface, this partial failure state cannot be replicated with one flexible supplier since it has only two states: both products fail with probability ρ or both products succeed with probability $(1-\rho)$. That is, a flexible supplier does not have any partial failure states. However, with product-dedicated suppliers, total failures (i.e., both suppliers fail) can still occur with probability ρ^2. Combining the above partial failure state with total failure state, the effective failure probability associated with product 1 is $\rho(1-\rho)+\rho^2=\rho$, which is exactly the failure probability for product 1 associated with the flexible supplier. Therefore, if embedding flexibility does not impact a supplier's risk profile (i.e., probability of failure), then the failure probability for each product remains intact. Consequently, embedding flexibility cannot impair the supply network's performance. Because it retains the same failure probability but brings additional demand-pooling benefit, embedding flexibility in a supplier is therefore always beneficial.

Note that a consequence of embedding flexibility in a supplier is the risk-aggregation effect; that is, the supply network experiences only extreme states: either total success with both products available or total failure with both products unavailable. These extreme states make the cash flows more volatile than what would have been observed in a supply network with product-dedicated suppliers. However, this volatility does not impact the expected cash flow, which is identical in both flexibility embedded supply networks and product-dedicated supply networks. For a risk-neutral firm, therefore, the risk-aggregation effect does not materialize as a financial disadvantage, which explains why embedding flexibility is always beneficial in the absence of reliable backup suppliers.

However, consider what happens when reliable backup suppliers are present. First, note that the preceding explanation (with no backup suppliers) hinges on the fact that partial failure states can be mapped into total failure states for each product. It turns out that this fact still holds in presence of reliable backup suppliers – as long as the backup suppliers are product-dedicated. A product-dedicated backup supplier will step in only when the primary supplier for that product fails. Therefore, although the backup suppliers interact with partial or total failure states, such interactions are product specific and are independent across products. For example, regardless of whether primary suppliers suffer from partial or total failures, the product-1 backup supplier's action is identical. That is, it will fill as much demand as possible for product 1. This action is independent of whether product 2 also failed and what actions were taken by the product-2 backup supplier. As such, partial failure states can still be mapped into total failure states for each product, implying that a flexibility-embedded supplier can still replicate the performance of product-dedicated suppliers. As a result, embedding flexibility into the supplier is again always beneficial even in presence of product-specific backup suppliers.

Interestingly, the above reasoning fails when the backup supplier is also flexible. A flexible backup supplier's action will be contingent upon whether primary suppliers suffer from partial or total failures because the backup supplier's capacity can be strategically allocated across the two products. As a result, a flexible backup supplier is utilized more efficiently in partial failure states than that in total failure states. Consequently, two implications arise. First, a flexible backup supplier is more valuable if primary suppliers suffer partial failures than when they suffer total failures. Second, a flexible *primary* supplier can no longer replicate product-dedicated primary suppliers since it does not have partial failure states, whereas a flexible backup supplier can be utilized more efficiently. Therefore, embedding flexibility into a primary supplier not only increases the volatility but also decreases the expected value of cash flow. In other words, the risk-aggregation effect of flexibility materializes as a negative financial impact on supply chain performance. When such negative impact becomes strong, the supply chain performs worse with a flexible primary supplier than that with product-dedicated suppliers. This means that the marginal benefit of adding more flexibility can be negative and can hurt supply chain performance. An important managerial implication is that embedding flexibility in a cheaper but unreliable primary supplier can be counterproductive if a flexible backup supplier is already in place.

The critical insight from this research is that it is the type of the backup supplier that fundamentally alters the value of flexibility. With product-dedicated backup suppliers, embedding flexibility in primary suppliers is always beneficial. In contrast, with a flexible backup supplier, embedding additional flexibility in primary suppliers is not necessarily beneficial. Supply chain performance can deteriorate with saturated flexibility, even if the flexibility is costless. We provide upper bounds on the demand-pooling and supplier diversification benefits associated with flexible and dedicated suppliers, respectively, and we show that adding flexibility to a primary supplier is unlikely to be beneficial when a backup supplier is flexible and only relatively expensive.

It is worth pointing out that the primary-backup supply network can also be construed as a contingent dual sourcing network, where the cost and reliability of the suppliers are such that the firm places regular orders with the cheaper but unreliable supplier while also placing contingent orders with the more expensive but reliable supplier. The key distinguishing factor between a primary supplier and a backup supplier is the supplier's reliability profile. From this angle, our analysis suggests that (a) flexibility should be prioritized to a more reliable supplier and (b) oftentimes full flexibility is neither necessary nor beneficial. A firm therefore needs to carefully evaluate its supply and demand risks. Depending on the magnitude of these risks, a partially flexible network can dominate a fully flexible network.

The rest of the paper is organized as follows. We review relevant literature in §2 and explain our model in §3. We analyze the merits of different flexibility configurations, identify key drivers of our findings, and develop the bounds of partially and fully flexible networks in §4, while we discuss the potential applications of our results for several different business settings in §5. We conclude our findings and discussions in §6. All proofs are relegated to the Appendix. An unabridged version that contains numerical experiments can be found at https://ssrn.com/abstract=3656803.

2. Literature review

This research is related to two streams of literature: uncertain supply and resource flexibility. The uncertain supply literature can be further divided into random yield, random lead time, and random capacity studies. While all three forms can result in supply disruptions, our research is more closely related to the Bernoulli type of random yield studies where the supply is either all or nothing (e.g., Parlar and Perry 1996; Gurler and Parlar 1997; Tomlin and Wang 2005; Tomlin 2006; Babich et al. 2007; Yang et al. 2009; Babich 2010; DeCroix 2013; Schmitt et al. 2015; Iancu et al. 2017; He et al. 2019). The focus of the above literature is on operational strategies to mitigate supply disruption risk, such as how to rank suppliers, how to allocate orders among available suppliers, how much inventory buffering to build into a supply chain, or how much backup capacity to reserve. A comprehensive review of the various strategies for mitigating supply disruption risk can be found in Snyder et al. (2016).

Several studies show how the use of backup suppliers is an important strategy for mitigating supply disruption risk. Chakraborty et al. (2016), for example, study the value of backup suppliers with demand uncertainty. Their study, however, does not consider contracting decisions related to the units of backup capacity to reserve. Instead, in their model, the backup supplier serves as a spot market with unlimited capacity such that a buyer can purchase any volume desired, but at a higher price compared to other suppliers that are subject to disruption. Zeng and Xia (2015) consider minimum purchase quantity required for maintaining good relationships with backup suppliers, but Sting and Huchzermeier (2010) point out that any firm (minimum) commitment contract imposes a dead weight loss on the system. Sting and Huchzermeier (2010) further analyze the effectiveness of different types of backup supplier contracts from a system efficiency perspective, and they find that capacity reservation contracts (which cover a given reservation quantity that can be delivered during a specific time window) are most effective. Rogers and White (2013) describe some key features of capacity reservation contracts: a) An upper limit on quantity delivery in a period; b) Buyer pays a price per unit for any portion of the reservation quantity that is not delivered in the period; and c) Buyer pays a higher price per unit for those units (up to reservation quantity) that are delivered in the period. In a similar vein, Wu et al. (2005) emphasize that the reservation fees, execution fees, and reservation quantity are common aspects specified in capacity reservation contracts. Our model setup adopts a similar contract structure to the structure outlined in Rogers and White (2013) and Wu et al. (2005).

Two commonly used alternative strategies for mitigating supply disruption risk are buffer inventory and dual sourcing. A buffer inventory approach is often applied to high-volume, low-value products. Simchi-Levi et al. (2015), however, show that Ford shunned buffer inventory due to the lean strategy the company was applying, despite the fact that many of its supply chain nodes subjected to high disruption risk were associated with low value parts (p. 381). Furthermore, an inventory buffer approach is not attractive in short life cycle industries where high obsolescence costs render inventory expensive, thus making the inventory buffer approach less desirable (Van Mieghem and Allon 2015). Moreover, buffer inventory can be stocked only if the original supply is not interrupted, implying that the target buffer inventory level may not be achievable at times, which hinders its ability to effectively mitigate disruption risk. For these reasons, our study does not consider this strategy.

Dual-sourcing or multiple-sourcing is another popular alternative to mitigate disruption risk. Two interesting and related challenges with this approach are (a) how to select a subset of suppliers from all available suppliers, and (b) how to allocate production orders among the selected suppliers. The first challenge can be addressed by finding effective measures to rank suppliers, such as by the cost/yield ratios (Federgruen and Yang 2009a,b). The second challenge is more difficult to overcome because allocating production orders based on ranking can often degenerate into a single sourcing policy when the highest ranked supplier has ample capacity. This result is not desirable from a risk mitigation perspective since the original purpose of multiple sourcing is to reduce supply disruption risk. In practice, managers often use ad hoc heuristics to allocate production orders to ensure that several suppliers are allocated some production orders, which reduces supply disruption risk and maintains on going relationships. The modeling framework of our study can be seen as a special case of dual sourcing, where backup suppliers can be viewed as contingent suppliers. A key difference between traditional dual-sourcing strategies and our approach, however, is that in our model backup suppliers always maintain a productive relationship with the firm by signing a contract ahead of production schedule so at a minimum they retain the capacity reservation fees. This contingent contracting arrangement avoids the above-mentioned shortcoming of the degenerative single sourcing scenario that can happen with a cost/yield ratio ranking approach.

Contrasting with the above supply risk literature, a main focus of the flexibility literature is on the value of product flexibility for mitigating demand risk (Jordan and Graves 1995, Harrison and Van Mieghem 1999, Netessine et al. 2002, Bish and Wang 2004, Van Mieghem 2007, Iravani et al. 2007, Pang and Whitt 2009, Ata and Van Mieghem 2009, Bassamboo et al. 2010, Jayaswala et al. 2011, Qi et al. 2015, Boyabatli et al. 2016). A flexible resource can be either non-exhaustible (e.g., a server with certain capacity) or exhaustible (e.g., a certain amount of inventory), but the majority of studies focus on non-exhaustible resources. Specifically, production capacity can be regarded as a non-exhaustible resource and an important decision is whether the capacity should be configured as a product-flexible or product-dedicated resource. These studies show that a product flexible resource can prioritize the production of higher margin products, which explains why using a flexible resource can lead to higher expected profit than using dedicated resources, other things being equal. In fact, Van Mieghem (1998) proves that product-flexible capacity is valuable for firms facing uncertain demands from two customer segments, even if the demands are perfectly positively correlated. An important insight from Van Mieghem (1998) is that flexibility is valuable not only in its ability to bring demand-pooling benefit, but also in its ability to take advantage of contribution-margin differences across different products. The flexibility literature typically assumes that (a) product-flexible capacity is more expensive than product-dedicated capacity and (b) both types of capacities are reliable. Consequently, a product-flexible facility naturally displaces multiple product-dedicated facilities if the cost is the same. This outcome can be seen as an operational benefit of product-flexible capacity since fewer facilities are required to meet diverse product demands.

Interestingly, the supply disruption literature typically does not examine how the mitigation strategies of supply risk interact with flexibility configurations to mitigate demand risk. Similarly, the above flexibility literature typically does not examine how the value of flexibility is impacted by the supply risk (besides the demand risk). Consequently, there have been few studies that examine the trade-off between supplier diversification benefit (in mitigating supply risk) and demand pooling benefit (in mitigating demand risk). Some notable exceptions include Tomlin and Wang (2005) and Schmitt et al. (2015). Tomlin and Wang (2005) consider both supply and demand risk; consistent with the above flexibility literature, they prove that for risk-neutral firms, a flexible resource is always preferred to multiple dedicated resources – even when both types of resources may subject to high failure probabilities. That is, regardless of suppliers' reliability, using a risky flexible supplier always leads to a higher expected profit than using multiple dedicated suppliers. Note that Tomlin and Wang (2005) do show that if a firm is risk averse, then using a risky flexible supplier can lead to worse performance than using multiple dedicated suppliers. Their study, however, does not consider reliable backup suppliers. Schmitt et al. (2015) contrast the expected inventory cost of a centralized (flexible) warehouse with that of multiple decentralized (dedicated) warehouses. They show that "the centralized system has a lower expected cost than the decentralized system in most of the instances we test, which confirms that, when demand is stochastic, the risk pooling effect usually holds even when supply is subject to disruption" (p. 205). Their study therefore corroborates Tomlin and Wang's (2005) findings that risk pooling dominates supply diversification. Also similar to Tomlin and Wang (2005), Schmitt et al. (2015) also exclude backup resources in their model. In contrast with these extant studies, we consider both primary and backup suppliers, but we do not consider risk attitude in this paper. We explore the supply diversification benefit of using multiple dedicated resources versus the demand pooling benefit of using a flexible resource when the supply network consists of both risky primary suppliers and reliable backup suppliers.

To sum up, the supply disruption literature has studied important mitigation strategies such as using buffer inventory, backup supply, and dual sourcing (supplier diversification), while the flexibility literature has extensively studied the benefits of resource flexibility. To the best of our knowledge, however, the extant literature has not addressed the emerging issues of how these mitigation strategies interact with flexibility configurations in a supply network that consists of both primary and backup suppliers. One key contribution of this research is to fill this gap and to provide insights on the optimal flexibility configuration of a basic supply network.

3. Model

Our model considers a firm that operates two product lines (labeled as 1 and 2). Demand for each product is stochastic and independent. (The effects of dependent, correlated demand on the risk pooling benefit are well understood [cf., Eppen 1979], so we exclude this factor from our analysis.) The firm has a primary supplier for each product line. The primary suppliers, however, may experience supply failures, hence the firm also contracts with backup suppliers to provide contingent capacity in case a primary supplier fails or when the primary supply is insufficient to meet demand. For the primary suppliers, the firm pays for only what is delivered. For the backup suppliers, the firm incurs an upfront reservation fee plus a unit exercise cost for actual capacity utilized.

There are four possible flexibility configurations of our supply network: a dedicated network, a partially flexible network with backup flexibility, a partially flexible network with primary flexibility, and a fully flexible network. A dedicated supply network (dd) consists of dedicated primary and backup suppliers for each product line. A partially flexible network can be achieved by embedding flexibility into the dedicated network by replacing existing dedicated suppliers with a flexible one or by developing one of the existing dedicated suppliers to be flexible and dropping the other one. When flexibility is embedded into a backup supplier, the result is a partially flexible network with backup flexibility (df); when flexibility is embedded into a primary supplier, the result is a partially flexible network with primary flexibility (fd). Finally, a fully flexible network (ff) can be achieved by embedding flexibility in both primary and backup suppliers. Figure 1 illustrates these four possible configurations.

In all of the above network configurations, each product line always has a primary supplier and a backup supplier. For the dedicated supply network (dd), the primary and backup suppliers are dedicated to their respective product lines. As flexibility is introduced to this dedicated network, the primary and/or the backup suppliers are shared across the product lines. We model these supply network configurations because they are applicable to many different settings, such as a firm sourcing two types of engines for two different product lines. Each engine can be sourced from

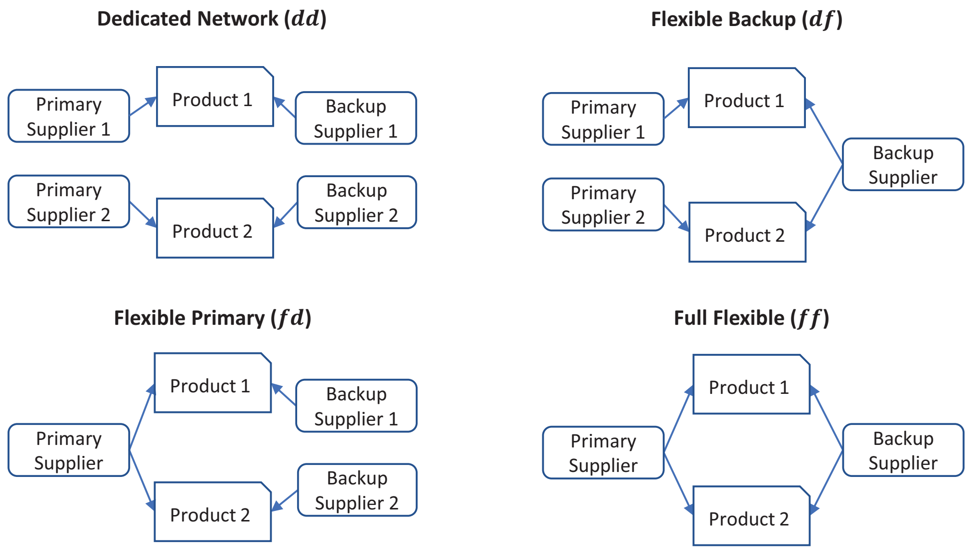

Figure 1 **Four possible configurations of a two-product supply network. The two-letter schematic identifies whether a primary or backup supplier is dedicated (d) or flexible (f), respectively.**

dedicated suppliers (i.e., using the dedicated supply network), or each engine can be sourced from a flexible supplier (i.e., using a variant of the flexible supply network depicted in Figure 1).

Next, we introduce necessary notations for the model. Without loss of generality, we normalize the unit costs of production for the two products to be equal and assume that product 1 is more profitable than product 2. When appropriate, a subscript j is appended to parameters for a specific product line. The following list outlines the notations for all network configurations.

- π_j: the unit contribution margin of product j, $j \in \{1, 2\}$
- s: the amount of backup capacity reserved
- q: the order quantity for the primary supplier (set before supply and demand risk is resolved)
- ρ: the probability that the primary supplier fails to deliver
- $\tilde{d}_j$: the stochastic demand for product j, $j \in \{1, 2\}$
- c: the unit production cost associated with the primary supplier
- γ: the unit upfront reservation cost associated with the backup supplier
- v: the unit exercise cost associated with the backup supplier

To isolate the value of flexibility along different network configurations, we assume that the cost and revenue parameters are independent of whether a dedicated or flexible supplier is used. While in reality a flexible supplier might be more expensive than a dedicated supplier, the implications of increased cost with a flexible supplier are obvious (all else being equal) and thus we do not account for them in the model. The revenue parameters, however, are distinguished between the two product lines. We also assume that the failure probability associated with a primary supplier is independent of whether the supplier is flexible or dedicated. Finally, we assume that $v \geq c$ so that primary production is prioritized when available. This implies that $\gamma + v > c$, which is reasonable as otherwise the primary supplier is never used.

Additionally, we assume the following sequence of events. First, before supply and demand risks are realized, regular orders are placed with primary suppliers and contingent capacity reservations are placed with backup suppliers. Second, supply status is resolved and the primary supplier either delivers the ordered quantity at a unit cost of c or fails to deliver. Third, demand is realized and the backup production is utilized (up to the reserved quantity) to satisfy remaining demand that cannot be met by deliveries from the primary suppliers. In essence, the primary suppliers operate under a make-to-stock mode while the backup suppliers operate under a make-to-order mode. This setting reflects the scenarios where the primary suppliers have a longer lead time (e.g., due to geographical locations in lower cost production regions) while the backup suppliers have a shorter lead time (e.g., located closer to the focal market with higher production cost) (Allon and Van Mieghem 2010). We now formulate the objective functions for each of the four network configurations.

3.1. Dedicated supply network (dd).

Let u_{dd} denote the expected profit associated with the dedicated network (dd). We have

$$\begin{aligned} u_{dd}(s_1, s_2, q_1, q_2) = &-\gamma(s_1+s_2) + \mathsf{E}\left[(\pi_1 - v)\min\left\{\tilde{d}_1, s_1\right\} + (\pi_2 - v)\min\left\{\tilde{d}_2, s_2\right\}\right]\rho^2 \\ &+ \mathsf{E}\left[\begin{array}{l} -cq_1 + \pi_1 \min\left\{\tilde{d}_1, q_1\right\} + (\pi_1 - v)\min\left\{\left(\tilde{d}_1 - q_1\right)^+, s_1\right\} \\ +(\pi_2 - v)\min\left\{\tilde{d}_2, s_2\right\} \end{array}\right](1-\rho)\rho \\ &+ \mathsf{E}\left[\begin{array}{l} -cq_2 + (\pi_1 - v)\min\left\{\tilde{d}_1, s_1\right\} + \pi_2 \min\left\{\tilde{d}_2, q_2\right\} \\ +(\pi_2 - v)\min\left\{\left(\tilde{d}_2 - q_2\right)^+, s_2\right\} \end{array}\right]\rho(1-\rho) \\ &+ \mathsf{E}\left[\begin{array}{l} -c(q_1+q_2) + \pi_1 \min\left\{\tilde{d}_1, q_1\right\} + (\pi_1 - v)\min\left\{\left(\tilde{d}_1 - q_1\right)^+, s_1\right\} \\ +\pi_2 \min\left\{\tilde{d}_2, q_2\right\} + (\pi_2 - v)\min\left\{\left(\tilde{d}_2 - q_2\right)^+, s_2\right\} \end{array}\right](1-\rho)^2, \end{aligned} \tag{1}$$

where $x^+ = \max(0, x)$. The first term is the reservation cost for backup capacity, the rest of the four terms together capture the expected profit under four possible delivery states for the primary suppliers: both fail to deliver with probability ρ^2, one succeeds and one fails with probability $\rho(1-\rho)$, and both succeed with probability $(1-\rho)^2$. It is straightforward to verify that (1) can be re-arranged as

$$u_{dd}(s_1, s_2, q_1, q_2) = -\gamma(s_1+s_2) + \mathsf{E}\left[(\pi_1 - v)\min\left\{\tilde{d}_1, s_1\right\} + (\pi_2 - v)\min\left\{\tilde{d}_2, s_2\right\}\right]\rho$$

$$+\left(-c(q_1+q_2)+\mathsf{E}\left[\begin{array}{c}\pi_1\min\left\{\tilde{d}_1,q_1\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+,s_1\right\}\\+\ \pi_2\min\left\{\tilde{d}_2,q_2\right\}+(\pi_2-v)\min\left\{\left(\tilde{d}_2-q_2\right)^+,s_2\right\}\end{array}\right]\right)(1-\rho). \tag{2}$$

This formula shows that all terms associated with $\rho(1-\rho)$ in the dd network can be subsumed into terms associated with ρ or $1-\rho$. That is, the original four possible delivery states for the primary suppliers can be reduced to two possible states: both suppliers fail to deliver or both succeed. We will subsequently show that this reduction of delivery states has important implications in the attractiveness of the dd network.

3.2. Flexible primary network (fd).

Let u_{fd} denote the expected profit associated with the flexible primary network (fd). We have

$$\begin{aligned}u_{fd}(s_1,s_2,q)=&-\gamma(s_1+s_2)+\mathsf{E}\left[(\pi_1-v)\min\left\{\tilde{d}_1,s_1\right\}+(\pi_2-v)\min\left\{\tilde{d}_2,s_2\right\}\right]\rho\\&+\mathsf{E}\left[\begin{array}{l}-cq+\pi_1\min\left\{\tilde{d}_1,q\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q\right)^+,s_1\right\}\\+\pi_2\min\left\{\tilde{d}_2,\left(q-\tilde{d}_1\right)^+\right\}+(\pi_2-v)\min\left\{\left(\tilde{d}_2-\left(q-\tilde{d}_1\right)^+\right)^+,s_2\right\}\end{array}\right](1-\rho).\end{aligned} \tag{3}$$

In contrast with the dd network above, there are only two possible cases for the primary supplier: success with probability $(1-\rho)$ or failure with probability ρ.

3.3. Flexible backup network (df).

Let u_{df} denote the expected profit associated with the flexible backup network (df). We have

$$\begin{aligned}u_{df}(s,q_1,q_2)=&-\gamma s+\mathsf{E}\left[(\pi_1-v)\min\left\{\tilde{d}_1,s\right\}+(\pi_2-v)\min\left\{\left(s-\tilde{d}_1\right)^+,\tilde{d}_2\right\}\right]\rho^2\\&+\mathsf{E}\left[\begin{array}{l}-c\left(q_1+q_2\right)+\pi_1\min\left\{\tilde{d}_1,q_1\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+,s\right\}\\+\pi_2\min\left\{\tilde{d}_2,q_2\right\}+(\pi_2-v)\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\end{array}\right](1-\rho)^2\\&+\mathsf{E}\left[\begin{array}{l}-cq_1+\pi_1\min\left\{\tilde{d}_1,q_1\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+,s\right\}\\+(\pi_2-v)\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\tilde{d}_2\right\}\end{array}\right]\rho\left(1-\rho\right)\\&+\mathsf{E}\left[\begin{array}{l}-cq_2+(\pi_1-v)\min\left\{\tilde{d}_1,s\right\}\\+\pi_2\min\left\{\tilde{d}_2,q_2\right\}+(\pi_2-v)\min\left\{\left(s-\tilde{d}_1\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\end{array}\right]\rho\left(1-\rho\right).\end{aligned} \tag{4}$$

Similar to the dedicated dd network, there are four possible cases for the primary suppliers. Unlike the dd network, however, the backup capacity is prioritized to the more profitable product line 1 in all four possible cases outlined above in (4).

3.4. Fully flexible network (ff).

Let u_{ff} denote the expected profit associated with fully flexible network (ff). We have

$$u_{ff}(s,q) = -\gamma s + \mathsf{E}\left[(\pi_1 - v)\min\left\{\tilde{d}_1, s\right\} + (\pi_2 - v)\min\left\{\left(s-\tilde{d}_1\right)^+, \tilde{d}_2\right\}\right]\rho$$
$$+\,\mathsf{E}\left[\begin{array}{l} -cq + \pi_1 \min\left\{\tilde{d}_1, q\right\} + (\pi_1 - v)\min\left\{\left(\tilde{d}_1 - q\right)^+, s\right\} \\ +\pi_2 \min\left\{\tilde{d}_2, \left(q-\tilde{d}_1\right)^+\right\} \\ +(\pi_2 - v)\min\left\{\left(s - \left(\tilde{d}_1 - q\right)^+\right)^+, \left(\tilde{d}_2 - \left(q - \tilde{d}_1\right)^+\right)^+\right\} \end{array}\right](1-\rho)\,. \qquad (5)$$

Similar to the flexible backup network *df*, there are only two possible cases for the primary supplier. Unlike the *df* network, however, the backup capacity is prioritized to the more profitable product.

4. Analysis of flexibility

Relative to the dedicated network *dd*, embedding flexibility in a backup supplier is always beneficial. This can be seen by the fact that, in a worst-case scenario, a flexible backup supplier can replicate dedicated backup suppliers by setting the reservation capacity level at $s = s_1 + s_2$. Therefore, all else being equal, the *df* network can perform no worse than the *dd* network. In addition, the added flexibility allows the backup capacity to be utilized more efficiently when primary suppliers fail. The following proposition follows from the above reasoning:

Proposition 1. *(a) All else being equal, embedding flexibility in a backup supplier is always beneficial. (b) The fully flexible network ff dominates the partially flexible primary network fd, i.e., $u_{ff} \geq u_{fd}$. (c) The partially flexible backup network df dominates the dedicated network dd, i.e., $u_{df} \geq u_{dd}$.*

Proposition 1 establishes a clear ranking of ff versus fd, as well as *df* versus *dd*. The relative performance between the partially flexible primary network fd and the dedicated network *dd*, as well as the relative performance between the fully flexible network ff and the partially flexible backup network *df*, however, are not immediately apparent, thus we explore them in the following subsections.

4.1. Flexible primary network (fd) versus dedicated supply network (dd)

The rationale that supports embedding flexibility in a backup supplier breaks down when it is applied to a primary supplier in an fd network. The reason is that primary suppliers can be unreliable, hence it is not immediately clear whether a single flexible supplier can replicate two dedicated suppliers: there are two possible supply states with a flexible supplier (either success or failure), but there are four possible supply states with the dedicated suppliers (both succeed;

both fail; supplier 1 succeeds, but supplier 2 fails; and supplier 1 fails, but supplier 2 succeeds). The additional supply states associated with dedicated primary suppliers potentially increase the *supplier diversification* benefit (less chance of complete supply failure), while a flexible primary supplier brings the *demand pooling* benefit (less chance of unsatisfied demand). In general, it is not obvious which benefit – supplier diversification or demand pooling – dominates in an arbitrary setting with both supply and demand risks. One might reason that the demand pooling benefit dominates when demand risk is much higher than supply risk, whereas the supply diversification benefit dominates otherwise. This reasoning, however, is incorrect in general. Tomlin and Wang (2005) establish that the demand pooling benefit always dominates the supplier diversification benefit for a risk-neutral firm, regardless of the supply failure probability or the magnitude of the demand risk. We build on Tomlin and Wang (2005) and provide an alternative proof that allows for arbitrary backup capacity and sheds light on why the supplier diversification benefit is dominated by the demand pooling benefit. Our proof generalizes the two-product result in Tomlin and Wang (2005) by allowing for positive backup capacities.

The key idea of our proof is to isolate the supplier diversification benefit and the demand pooling benefit associated with dd and fd networks, respectively. More concretely, we aim to relate the dd and fd networks with the following structure:

$$u_{dd} = u_{fd} + \mathrm{SD} - \mathrm{DP},$$

where SD denotes supplier diversification benefit whereas DP denotes the demand pooling benefit. Observe that the SD benefit occurs only when partial supply failures occur (i.e., one supplier succeeds while the other fails). Thus, the SD benefit must be comprised of terms associated with $\rho(1-\rho)$. Interestingly, while the dd network does have terms associated with $\rho(1-\rho)$ (see (1)), some re-arrangements of terms show that all terms associated with $\rho(1-\rho)$ can be subsumed into terms associated with ρ or $1-\rho$ (see (2)). It directly follows that the SD benefit associated with the dd network must be zero (SD $= 0$). In other words, the supplier diversification benefit (relative to the flexible primary network fd) does not materialize.

On the other hand, the DP benefit occurs only when the flexible primary supplier succeeds in delivery. Thus, the DP benefit must be comprised of terms associated with $1-\rho$. Given the fact that SD $= 0$, the DP benefit must be captured entirely by the revenue difference between the dd and fd networks. To be more concrete, let $q = q_1 + q_2$ and substitute (3) into (2), we have

$$u_{dd}(s_1, s_2, q_1, q_2) = u_{fd}(s_1, s_2, q_1 + q_2) - \mathrm{DP}(s_1, s_2, q_1, q_2),$$

where

$$\mathrm{DP}(s_1,s_2,q_1,q_2)=\mathsf{E}\left[\begin{array}{l} \left(\begin{array}{l} \quad \pi_1 \min\left\{\tilde{d}_1, q_1+q_2\right\}+\pi_2 \min\left\{\tilde{d}_2, \left(q_1+q_2-\tilde{d}_1\right)^+\right\} \\ +\ (\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1-q_2\right)^+, s_1\right\} \\ +\ (\pi_2-v)\min\left\{\left(\tilde{d}_2-\left(q_1+q_2-\tilde{d}_1\right)^+\right)^+, s_2\right\} \end{array}\right) \\ -\left(\begin{array}{l} \quad \pi_1 \min\left\{\tilde{d}_1, q_1\right\}+\pi_2 \min\left\{\tilde{d}_2, q_2\right\} \\ +\ (\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+, s_1\right\} \\ +\ (\pi_2-v)\min\left\{\left(\tilde{d}_2-q_2\right)^+, s_2\right\} \end{array}\right) \end{array}\right](1-\rho)\geq 0.$$

In the above expression, the top terms capture the sales from primary/backup production under ff and the bottom terms under df. Their difference reflects the demand pooling benefit when the primary suppliers are up. Because $u_{fd}(s_1,s_2,q^*)\geq u_{fd}(s_1,s_2,q_1+q_2)$, it follows that $u_{fd}(s_1,s_2,q^*)\geq u_{fd}(s_1,s_2,q_1+q_2)\geq u_{dd}(s_1,s_2,q_1,q_2)$ (i.e., the fd network dominates the dd network). Notice that the above derivation does not depend on any specific supply and demand risk profiles or backup capacity levels. Therefore, the result holds for any arbitrary supply or demand risk profiles, as well as for any arbitrary backup capacity levels.

The two-product findings for a risk-neutral firm in Tomlin and Wang (2005) can be viewed as a special case of our model by setting $s_1=s_2=0$. For this special case, the demand pooling benefit is simplified to

$$\mathrm{DP}(0,0,q_1,q_2)=\mathsf{E}\left[\begin{array}{l} \quad \pi_1 \min\left\{\tilde{d}_1, q_1+q_2\right\}+\pi_2 \min\left\{\tilde{d}_2, \left(q_1+q_2-\tilde{d}_1\right)^+\right\} \\ -\ \pi_1 \min\left\{\tilde{d}_1, q_1\right\}-\pi_2 \min\left\{\tilde{d}_2, q_2\right\} \end{array}\right](1-\rho)\geq 0. \tag{6}$$

The above analysis shows that the demand pooling benefit is present in the fd network but the supplier diversification benefit is absent in the dd network. The following theorem gives key insights into the supplier diversification benefit.

THEOREM 1. *Supplier diversification conferred by a dedicated supply network dd offers no value if partial failure states $\rho(1-\rho)$ can be reduced to total success state $1-\rho$ or total failure state ρ.*

The fact that supplier diversification offers no value for a risk-neutral firm is not immediately obvious ex ante. Supplier diversification reduces the chance of total supply failures, thus intuitively, it should be valuable. However, as our findings show, supplier diversification is a necessary but not sufficient condition for SD benefit to accrue. The *dd* network creates the opportunity for the supplier diversification benefit, but it does not provide the mechanism to capture that benefit. Any mechanism that captures the potential value of partial failures must interact with partial failure states $\rho(1-\rho)$ so that they cannot be reduced to total success or failure states. The dd network lacks any such mechanism, hence its supplier diversification benefit does not materialize.

Because the demand pooling benefit is positive (i.e., DP ≥ 0 [see (6)]), the fd network dominates the dd network (i.e., $u_{fd} \geq u_{dd}$). By Proposition 1(b), $u_{ff} \geq u_{fd}$. Combining these results yields $u_{ff} \geq u_{fd} \geq u_{dd}$. This inequality shows that as flexibility progressively replaces dedicated resources in dd, starting with the primary supplier and moving to the backup supplier, the supply network performance steadily improves. At the same time, Proposition 1(c) indicates that $u_{df} \geq u_{dd}$; that is, adding partial flexibility to a dd network by replacing two dedicated backup suppliers with one flexible supplier alone also improves the supply network performance. This suggests that a firm has a range of latitude in incorporating flexibility in its supply network and can be assured that the embedding flexibility in either primary or backup suppliers is advantageous. However, the above analysis does not speak to the relative performance between the partially flexible backup network df and the fully flexible network ff. We explore their relative performance in the next subsection.

4.2. Fully flexible network (ff) versus flexible backup network (df)

When comparing a fully flexible supply network ff with a partially flexible backup network df, one might intuitively believe that the ff network must dominate the df network. The reason for this intuition seems straightforward: all else being equal, a fully flexible network can never do worse than a partially flexible network as long as the unit production costs of a flexible supplier are on par with those of a product-dedicated supplier. The above reasoning, however, turns out to be incorrect: an ff network may perform worse than a df network – even if flexibility is costless.

To understand this finding, consider a simple setting where demand for each product is fixed at d (i.e., $\tilde{d}_1 = \tilde{d}_2 = d$). With such a setting in mind, the following proposition illustrates the performance of a df network versus an ff network.

PROPOSITION 2. *If demand is deterministic, then df dominates ff. Furthermore, this dominance is strict whenever $\rho(\pi_2 - v) < \gamma < \rho(\pi_1 - v)$.*

Proposition 2 shows that a partially flexible network can dominate a fully flexible network. The intuition for this result is that, with deterministic demand, the DP benefit disappears but the SD benefit persists. It is intuitive that DP=0 under deterministic demand, but not apparent why SD can be positive. The reason is that the df network utilizes the backup supplier more efficiently under partial failures (i.e., more likely to cover demand for both products). In contrast, the ff network does not use the backup supplier when the flexible primary supplier succeeds (so DP=0) but relies on the backup supplier entirely when the primary supplier fails (which cannot cover both products' demand unless s is sufficiently large). An acute reader, however, may immediately recognize that such an example is trivial because it is not meaningful for a firm to embed flexibility in primary suppliers when demand is deterministic. Nevertheless, the only reason that the partially flexible network can dominate must be due to the supplier diversification benefit. The presence of the supplier diversification benefit in a deterministic demand setting hints that *df* may dominate ff even in more general settings with stochastic demand.

Suppose now we partially relax the above deterministic setting by allowing the demand for product 2 to be stochastic while keeping the demand for product 1 deterministic (i.e., $\tilde{d}_1 = d$ and $\mathsf{E}[\tilde{d}_2] = d$). Let G denote the distribution of $\tilde{d}_2$. Stochastic demand for product 2 allows the opportunity for demand pooling benefit to arise with ff network but not with the *df* network. In this setting, it is not immediately obvious whether *df* always dominates ff or vice versa.

PROPOSITION 3. *If demand is deterministic for product 1 but stochastic for product 2 (i.e., $\tilde{d}_1 = d$ and $\mathsf{E}[\tilde{d}_2] = d$), then (a) df always dominates ff, (b) this dominance is strict whenever $s^{ff} \in (0, d]$ and $s^{ff} < G^{-1}\left(\frac{\pi_2 - c}{\pi_2 - v}\right)$, and (c) demand uncertainty reduces the degree of df dominance over ff.*

In this partially stochastic demand setting, the fully flexible network ff never dominates the partially flexible network *df*. This result is somewhat perplexing, as one would have expected that DP should offset SD at least sometimes. It turns out that the DP benefit never materializes, but the SD benefit persists. The reason is that, regardless of the magnitude of product 2 demand, it is always optimal under ff to allocate d units to product 1 and the rest to product 2. Thus, there is no opportunity for the ff network to adjust allocations based on realized product 2 demand. In contrast, the SD benefit persists because partial failures allow a more efficient use of the backup supplier. Nevertheless, part (c) suggests that demand uncertainty reduces the SD benefit. The intuition is that large swings in product 2 demand make it more difficult to cover both products' demand with partial failures, thus reducing the value of SD. Hence, while the above offers convincing evidence that the SD benefit dominates, it is not clear whether or when such a dominance may continue to hold for general stochastic demand.

We now generalize the above analysis to characterize the relative performance between ff and *df*. We establish a high-level result that helps explain the supplier diversification benefit.

PROPOSITION 4. *The partial supply failure states associated with df cannot be reduced to total success or failure states when backup capacity $0 < s < \max\left\{\tilde{d}_1 + \tilde{d}_2\right\}$.*

Note that if $s = 0$, then backup allocation is irrelevant and the two partial failure states can be decoupled into total success and failure states for each specific product. On the other hand, if $s > \max\left\{\tilde{d}_1 + \tilde{d}_2\right\} \Rightarrow q_1 = q_2 = 0$, then the two partial failure states can again be decoupled into total success or failure states for each specific product. This observation leads to the following theorem.

Theorem 2. *Irreducible partial failure states are a necessary and sufficient condition for the supplier diversification benefit to exist.*

We are now ready to characterize the supplier diversification benefit and the demand pooling benefit for *df* and ff networks. Recall that the supplier diversification benefit must be related to partial failures (i.e., terms associated with $\rho(1-\rho)$), while the demand pooling benefit can only occur when the flexible primary supplier succeeds (i.e., terms associated with $(1-\rho)$). Let $q_1+q_2=q$, then $u_{df}(s,q_1,q_2)=u_{ff}(s,q_1+q_2)+\mathrm{SD}(s,q_1,q_2)-\mathrm{DP}(s,q_1,q_2)$, where

$$\mathrm{DP}(s,q_1,q_2)=\mathsf{E}\left[\sum_{i=1}^{2}\left\{\pi_i(x_i^{ff}-x_i^{df})+(\pi_i-v)\,(y_i^{ff}-y_i^{df})\right\}\right](1-\rho),$$
$$\mathrm{SD}(s,q_1,q_2)=\mathsf{E}\left[(\pi_2-v)\left\{(z_{10}-z_{11})+(z_{01}-z_{00})\right\}\right]\rho(1-\rho).$$

The x_i^k and y_i^k are primary and backup sales of product i under configuration $k\in\{ff,df\}$, respectively. Hence, $\mathrm{DP}(s,q_1,q_2)$ captures the sales difference between ff and *df* when the primary production is operational (i.e., the demand pooling benefit). On the other hand, z_{ij} are backup production volumes of product 2 under *df*, where the subscripts denote whether the primary suppliers for product 1 and 2 is up (1) or down (0). Hence, $\mathrm{SD}(s,q_1,q_2)$ captures the difference in backup production volumes when one of the primary suppliers fails versus when both fail or both succeed (i.e., the supplier diversification benefit). The detailed expressions for x_i^k, y_i^k, and z_{ij} can be found in the proof of Proposition 5.

Proposition 5. *(a)* $\mathrm{DP}\geq 0$. *(b)* $\mathrm{SD}\geq 0$.

The result that $\mathrm{DP}\geq 0$ is intuitive because (a) there are no instances of excess primary production of product 1 with simultaneous shortage of primary production of product 2 and (b) primary production capacity is allocated to the most profitable product (e.g., the constraint requiring that q_i must be allocated to product i demand is relaxed). The reason for $\mathrm{SD}\geq 0$ is more nuanced. Part of the intuition is that (a) the backup production is the same (between *df* and ff) when primary suppliers are both operational or down, and (b) the reserved backup capacity is utilized more efficiently under partial failures (i.e., less chance of shortage or excess in backup capacity reservation).

Proposition 6. *(a)* $\mathrm{DP}=0$ *if product 1 demand is deterministic. (b)* $\mathrm{SD}=0$ *if* $s=0$ *or* $s=\max\left\{\tilde{d}_1+\tilde{d}_2\right\}$.

Proposition 6 gives conditions under which either demand pooling benefit or supplier diversification benefit does not materialize. The key factor that enables the demand pooling benefit is the uncertainty associated with a *more profitable* product, while the key factor that enables the supplier diversification benefit is a moderate amount of *flexible* backup capacity.

The partially stochastic demand setting analyzed earlier corresponds to case (a) of Proposition 6. In this setting, because product 1 demand is deterministic, the demand pooling benefit is $\mathrm{DP}=0$. The supplier diversification benefit, however, is robust against deterministic product 1 demand. This result can be verified by setting the primary production for product 1 equal to the deterministic demand (i.e., $q_1=d$).

$$\mathrm{SD}(s,d,q_2)=(\pi_2-v)\,\mathsf{E}\left[\begin{array}{l}\left(\min\left\{s,\tilde{d}_2\right\}-\min\left\{s,\left(\tilde{d}_2-q_2\right)^+\right\}\right)\\-\left(\min\left\{(s-d)^+,\tilde{d}_2\right\}-\min\left\{(s-d)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\right)\end{array}\right]\rho(1-\rho)\geq 0.$$

Because $\mathrm{SD}\geq 0$ and $\mathrm{DP}=0$ in the partially stochastic demand setting, the partially flexible network df dominates the fully flexible network ff.

The two-product case for a risk neutral firm in Tomlin and Wang (2005) corresponds to case (b) of Proposition 6. In their model, because the backup capacity is zero, the supplier diversification benefit is $\mathrm{SD}=0$. The demand pooling benefit, however, is robust against the amount of backup capacity. This result can be verified by setting the backup capacity to zero (i.e., $s=0$).

$$\mathrm{DP}(0,q_1,q_2)=\mathsf{E}\left[\begin{array}{l}\left(\pi_1\min\left\{\tilde{d}_1,q_1+q_2\right\}+\pi_2\min\left\{\tilde{d}_2,\left(q_1+q_2-\tilde{d}_1\right)^+\right\}\right)\\-\left(\pi_1\min\left\{\tilde{d}_1,q_1\right\}+\pi_2\min\left\{\tilde{d}_2,q_2\right\}\right)\end{array}\right](1-\rho)\geq 0.$$

Because $\mathrm{DP}\geq 0$ and $\mathrm{SD}=0$ in the no-backup setting, the fully flexible network ff dominates the partially flexible network df.

In a general setting with stochastic demand and positive backup capacity, both the supplier diversification benefit associated with df and the demand pooling benefit associated with ff can be positive. Depending on the magnitudes of these two benefits, the partially flexible network df may dominate the fully flexible network ff. That is, targeting a fully flexible network may not be desirable: it may perform worse than a partially flexible network.

4.3. When costless flexibility hurts

Flexibility hurts whenever the supplier diversification benefit of df dominates the demand pooling benefit of ff. While the supplier diversification benefit SD and the demand pooling benefit DP can be easily computed numerically, it is of interest to gain more transparent understanding of when SD is likely to dominate DP. Proposition 6 suggests that SD is closely related to the backup supplier. A key intuition is that the df network utilizes the backup capacity s more efficiently than the ff network does. Thus the SD benefit is more likely to dominate when the system parameters favor more efficient utilization of the backup capacity under df. Guided by this intuition, the following proposition characterizes a general condition that ensures df dominance.

Proposition 7. *Suppose $v < \pi_2$. For any given backup capacity s, the df network dominates the ff network if*

$$\rho\mathcal{H}(s) \geq \frac{c}{\pi_2 - v}, \tag{7}$$

where

$$\mathcal{H}(s) = \left(\frac{1}{z}\right) \mathsf{E}\left[\left(\begin{array}{c} \min\{(s-(\tilde{d}_1 - q^{ff})^+)^+, \tilde{d}_2\} \\ -\min\{(s-(\tilde{d}_1 - q^{ff})^+)^+, (\tilde{d}_2 - z)^+\} \end{array}\right) - \left(\begin{array}{c} \min\{(s-\tilde{d}_1)^+, \tilde{d}_2\} \\ -\min\{(s-\tilde{d}_1)^+, (\tilde{d}_2 - z)^+\} \end{array}\right)\right],$$

q^{ff} is the optimal primary production quantity under ff, and z is the unique solution to

$$\pi_2 \mathsf{E}\left[\begin{array}{c} \min\{\tilde{d}_2, z\} \\ -\min\{\tilde{d}_2, (q^{ff} - \tilde{d}_1)^+\} \end{array}\right] = (\pi_2 - v)\mathsf{E}\left[\begin{array}{c} \min\{(s-(\tilde{d}_1 - q^{ff})^+)^+, (\tilde{d}_2 - (q - \tilde{d}_1)^+)^+\} \\ -\min\{(s-(\tilde{d}_1 - q^{ff})^+)^+, (\tilde{d}_2 - z)^+\} \end{array}\right].$$

Further, the condition defined by $\rho\mathcal{H}(s) \geq c/(\pi_2 - v)$ is non-empty, that is, there exist parameter values that guarantee the above condition is satisfied such that the df network dominates the ff network.

Condition (7) is sufficient, but not necessary, for *df* dominance because *df* production quantities are derived from the optimal ff solution (i.e., $q_2 = z, q_1 = q^{ff} - z$). Function $z\mathcal{H}(s)$ is structurally similar to the expected quantity in SD; so the expression captures the increase in sales volume of *df* that is associated with backup supply. Accordingly, $\mathcal{H}(s)$ can be interpreted as the fraction of stocking units intended for product 2 that is actually sold – facilitated by the backup capacity s. This interpretation implies that if the pre-positioned primary supply for product 2 is more likely to be sold, then the *df* network is more likely to dominate the ff network. While this result might appear obvious ex-post, it is not apparent ex ante, especially considering the fact that the above condition must be facilitated by the backup capacity. Notice that $\mathcal{H}(0) = 0$ and $\mathcal{H}(\infty) = 0$, corroborating our earlier result that for extreme values of s, the *df* network cannot dominate the ff network.

The above result confirms that the *df* network can dominate the ff network for general stochastic demand. Furthermore, condition (7) suggests that, all else being equal, the *df* network is more likely to dominate the ff network when ρ is not too small, v is moderately large (but must be less than π_2), and c is small. Hence, the *df* network is likely to be more attractive when the primary suppliers are risky but inexpensive. However, condition (7) involves function $\mathcal{H}(s)$, which itself is also influenced by ρ, v, and c implicitly through q^{ff}. We therefore caution that the above interpretation only reflects some first-order effects. The following proposition gives a more transparent characterization for *df* dominance.

Proposition 8. *Suppose $v < \pi_2$ and $0 < \rho < 1$. (a) There exists a unique value of Γ such that for some $\gamma + v \le \Gamma$, the df network (weakly) dominates the ff network. Conversely, for some $\gamma + v > \Gamma$, the ff network can strictly dominate the df network. (b) Furthermore, $\partial\Gamma/\partial c \le 0$.*

Part (a) of Proposition 8 suggests that the *df* network is more likely to dominate when the backup capacity is not too expensive ($\gamma + v$ less than some threshold). This result corroborates part (b) of Proposition 6, which states that the supplier diversification benefit can only exist with some amount of backup capacity. For the supplier diversification benefit to overcome the demand pooling benefit, the effective cost of the backup capacity $\gamma + v$ must not exceed a certain threshold. Note that when $\gamma + v \to 0$, the performance of *df* and ff converge (i.e., *df* only weakly dominates).

Proposition 8(b) tells us that the primary supplier cost plays an important role in *df* dominance: as the primary supplier cost increases, the backup supplier cost must decrease to maintain *df* dominance. This echoes Proposition 7, which states that the *df* network is more likely to dominate when c is small (i.e., when primary suppliers are inexpensive). The intuition is that when the primary supplier is expensive, the optimal primary production quantity decreases, which amplifies the demand pooling benefit (since ff can utilize limited supply more efficiently) but reduces supplier diversification benefit (since *df* must rely more heavily on the backup supply to cover demand when partial failure occurs).

The above results can be summarized as follows: as long as (a) primary suppliers are unreliable ($0 < \rho < 1$) and (b) backup capacity is viable ($v < \pi_2$), there *always* exists an effective threshold cost associated with the backup supply such that *df* (weakly) dominates. However, such a threshold is increasingly difficult to satisfy as the primary suppliers become more expensive – since backup supply is typically more expensive than primary supply. An important managerial implication is that in a typical sourcing setting where the primary suppliers are unreliable but inexpensive and the backup suppliers are reliable but more expensive, retaining dedicated primary suppliers may offer sufficient supplier diversification benefit, offsetting the demand pooling benefit offered by a flexible primary supplier. Nonetheless, such a trade-off between supplier diversification and demand pooling benefit must be weighed against other relevant factors, such as differences in fixed and variable costs for dedicated and primary suppliers, as well as other non-cost factors such as human resource and supply network implications.

To provide additional clarity on *df* or ff dominance, next we leverage the general results above to develop concrete, easy-to-interpret bounds for SD and DP benefits.

Proposition 9. *For any given s and $q = q_1 + q_2$, let $\overline{\mathrm{DP}}$ and $\overline{\mathrm{SD}}$ denote the upper bound of demand pooling and supplier diversification benefit, respectively. Then*

$$\overline{\mathrm{DP}} = \underbrace{(1-\rho)}_{(a)} \times \underbrace{\pi_1 q_2}_{(b)} \times \underbrace{\overline{G}_1(q_1)}_{(c)}, \quad \overline{\mathrm{SD}} = \underbrace{\rho(1-\rho)}_{(a)} \times \underbrace{(\pi_2 - v)s}_{(b)} \times \underbrace{G_1(q_1)}_{(c)}.$$

To provide concrete managerial guidance, we now interpret the above two bounds. For the demand pooling benefit, $\overline{\mathrm{DP}}$:

(a) increases in total success state probability $(1-\rho)$,

(b) increases in the uncertainty of product 1 demand, up to $(1-\rho)\pi_1 q_2 \geq 0$,

(c) tends to zero as the variance of product 1 demand approaches to zero.

For the supplier diversification benefit, $\overline{\mathrm{SD}}$:

(a) increases in partial failure states probability $\rho(1-\rho)$,

(b) decreases in the uncertainty of product 1 demand, and

(c) tends to $\rho(1-\rho)(\pi_2 - v)s \geq 0$ as the variance of product 1 demand approaches to zero.

The demand pooling benefit DP arises when realized demands for product 1 and 2 deviates from their critical newsvendor ratio so that the primary production can be dynamically adjusted between the two products. The maximum benefit is obtained when realized demand for product 1 is high, hence the margin $\pi_1 - c$ is an important driver. More importantly, the uncertainty of product 1 demand is a critical driver. If there is little uncertainty, then the primary production will match demand closely in the first place, leaving little opportunity for substitution. The maximum ceiling expression $(1-\rho)(\pi_1 - c)q_2$ aptly captures the above essential drivers.

The supplier diversification benefit SD arises when the partial failure states can be leveraged by the available backup capacity. To see this result, consider the four possible states for the primary suppliers. If both primary suppliers succeed, the backup capacity is unlikely to be utilized under either *df* or ff, hence no diversification benefit can accrue. If both primary suppliers fail, the backup capacity is likely insufficient to cover combined product demand under either *df* or ff, hence again, no diversification benefit can accrue. In contrast, if one of the primary suppliers fails but the other succeeds, then the backup capacity is likely to have leftovers after covering product 1 demand, which can be used to cover product 2 demand. This scenario can only happen with *df* and not with ff, hence the supplier diversification benefit only arises in these partial failure states. The marginal benefit obtained for each unit of backup capacity allocated to product 2 demand is $(\pi_2 - v)$ and the maximum possible units that can be allocated is s. As such, the maximum ceiling expression $\rho(1-\rho)(\pi_2 - v)s$ aptly captures the above drivers.

In summary, a partially flexible network *df* can dominate a fully flexible network ff. Such dominance is more likely to occur when primary suppliers are risky but inexpensive and while backup suppliers are reliable but moderately expensive. Although demand uncertainty amplifies the value of DP and dampens the value of SD, there always exists some threshold value for the effective cost (i.e., $\gamma + v$) of the backup capacity such that the SD benefit can offset the DP benefit. However, as the primary suppliers become more expensive, it is more challenging for the SD benefit to dominate, hence the ff network is more likely to be appealing. Firms, therefore, should carefully consider possible flexibility configurations for their supply networks when both supply and demand risks are significant. Unless the primary suppliers are very expensive, a partially flexible network with backup flexibility appears to be a robust first step. Consequently, depending on the market and cost structure, it can be advantageous not to implement fully flexible supply network.

5. Managerial implications: Extensions to other types of supply networks

While our model setup is centered around a product-specific and product-flexible supply networks, the insights generated from our study apply to other types of supply networks as well. Below, we illustrate a number of such extensions. It is worth emphasizing that no supplier is perfectly reliable; however, it is the relative difference in risk and cost profiles – not the absolute reliability of backup suppliers – that drive the key results. As such, knowing the relative difference between cost and reliability of suppliers is adequate for applying our insights to various types of networks.

5.1. Task-specific supply networks

Our study is applicable for task-specific supply networks, such as service systems where managers move up through the ranks acquiring skills in a variety of tasks. Such service systems can be call centers, automobile repair centers, or grocery stores. In these types of systems, there are usually two types of employees: those on hourly pay and those on salaried pay. In general, those on hourly pay tend to be specialized and less reliable (i.e., some may not show up on time for a variety of reasons). In this case, managers (salaried) are more reliable and can serve as contingent flexible backup supply if not enough hourly personnel show up.

In such service systems, an important management task is to determine whether or not to cross-train the hourly personnel on different tasks so as to reduce the total number of hourly personnel on the payroll. The consequence of not cross-training hourly employees is a system that relies on dedicated primary suppliers (hourly employees) with a flexible backup supplier (salaried managers). While there is a large stream of literature that advocates cross-training in many different settings, our analysis suggests that cross-training hourly personnel (with an intention to reduce payroll cost) may not be a good strategy since hourly workers are risky (i.e., they have a higher likelihood of churning and rotation). If there are sufficient salaried personnel who have worked up through the ranks and are familiar with the variety of tasks performed by hourly personnel, then cross-training may not be as attractive as staffing more hourly personnel. In other words, staffing a large number of dedicated hourly personnel can be more attractive than staffing a small number of flexible hourly personnel when salaried managers can serve as flexible backups.

5.2. Region-specific distribution networks

We also find applicability for region-specific distribution networks, such as regional distribution centers (RDCs) and central distribution centers (CDCs). RDCs act as primary suppliers that meet regional demand very quickly by serving a specific geographic region with a limited amount of inventory. If we interpret shipping time as cost and stock level as service level, then RDCs can be regarded as less reliable, lower-cost primary suppliers. Relative to CDCs, RDCs are less reliable because they have limited stock and therefore are more likely to experience stock outs; at the same time, they are cheaper because they can serve regional demand faster.

CDCs, on the other hand, typically have much higher inventory and service levels than RDCs, thus they are more reliable than RDCs. Additionally, CDCs can serve a specific geographic region of customers and can, at the same time, restock RDCs. Thus, CDCs can be used as reliable and flexible backup suppliers, shipping orders to customers when an RDC is out-of-stock. This option is more expensive than shipping directly from an RDC (e.g., it incurs longer shipping time), but the backup is more reliable RDCs. From a network design perspective, deciding the geographical region covered by an RDC is equivalent to deciding the level of primary supplier flexibility. In other words, increasing the geographical region covered by an RDC implies that a smaller number of RDCs are deployed in the network, which is equivalent to increasing the level of primary supplier flexibility (i.e., the RDCs can serve larger regions). In this case, adding flexibility (by having a smaller number of RDCs each covering a larger area) is not necessarily free: while a firm may save on fixed costs by maintaining fewer RDCs, the total cost may increase (i.e., delivery speed decreases). Nevertheless, our results suggest that even if a smaller number of RDCs can still maintain the same delivery speed (through automation and better logistic management, for example), the network performance can deteriorate (relative to a larger number of RDCs each covering a smaller area) – if the RDCs are relatively small (and thus less reliable than CDCs) and if CDCs can adequately serve as flexible backup suppliers for RDCs.

5.3. Manufacturing facility networks

Our model also helps large organizations that operate a supply network with facilities located in both low-cost and high-cost countries. The typical trade off these organizations face when designing a supply network to serve demand in high-cost countries is cost versus reliability. A facility in a low-cost country may offer a cost advantage, but due to greater distance from the high-cost country market, it may be more susceptible to replenishment lead time delays (e.g., delays caused by natural disasters, like hurricanes, or human-caused setbacks, like labor strikes at ports of exit or entry). A firm may set up facilities in the low-cost country to serve as primary suppliers and also set up facilities in the high-cost country to serve as backup suppliers to satisfy demand that may otherwise be lost during the delay. The results of our study can help organizations determine the merits of embedding flexibility in this type of supply networks.

From a network design perspective, configuring facility flexibility is an important strategic decision. While it might appear intuitive that embedding flexibility in low-cost country facilities (in order to reduce the number of facilities managed) is more beneficial, our study suggests that this approach may not be the best strategy when flexibility leads to a reduction in the number and/or geographic diversity of low-cost suppliers. When replenishment lead times from low-cost countries are highly uncertain, a flexible high-cost facility network (with dedicated low-cost facilities) can outperform a flexible low-cost facility network (with or without dedicated high-cost facilities). Furthermore, if a facility in a high-cost country is already flexible, then introducing additional flexibility to facilities in low-cost countries can be counterproductive.

5.4. Commonality design in spare-parts networks

This model also has important implications for businesses in spare-parts network industries, such as transportation, where maintenance and repair operations (MRO) comprise a large fraction of business spending. Spare parts are mission critical for MROs, and a key challenge is determining how spare parts can satisfy maintenance and repair needs across a variety of equipment. Spare parts can be designed and manufactured over a common platform so they can fit on different equipment, or they can be designed and manufactured specifically for each type of equipment. Using a common platform to produce universal spare parts is a flexible strategy that accommodates demand risks, but this strategy likely incurs a higher cost because the parts need to be able to function across different equipment. In contrast, individually designed parts have no flexibility in meeting demand since they can only fit on a particular type of equipment, but the manufacturing cost of this strategy can be cheaper due to the part's simpler design.

Our analysis indicates that it can be beneficial for a firm not to adopt commonality across its parts manufacturing network – when the network comprises both low-cost and high-cost facilities that have different risk profiles due to differences in physical distance, infrastructure, and management proximity. In some cases, assigning some low-cost facilities to manufacture dedicated parts while assigning a high cost facility to manufacture universal parts can be more advantageous than adopting commonality across the entire network with a reduced number of low-cost facilities. When only universal parts are manufactured, the cost savings associated with dedicated parts are lost. When lost savings are coupled with the supply risk associated with low-cost facilities, the demand pooling benefit of commonality can be wiped out. It is worth pointing out a related concern about a pure commonality approach: this approach may lead to system-wide failures if defects exist in design or manufacturing processes, so maintaining some dedicated parts facilities can reduce the chance of a system-wide failure. While our model does not consider design defects, the above concern is another factor that can reduce the attractiveness of commonality design.

5.5. Mission-critical, low latency networks

Lastly, our model has important managerial implications for the unconventional but emerging area of edge computing networks. Edge computing allows localized processing, which reduces latency but edge devices have a higher risk in dependability and capacity because "these edge devices are often not as well maintained, dependable, powerful, or robust as centralized server-class cloud resources" (Bagchi et al. 2020). In such systems, having a powerful cloud-based resource is necessary for maintaining the continued operations of the network. Applying our model to this type of network, edge computing devices can be viewed as low-cost, dedicated suppliers within a restricted locale, while the cloud-based resource serves as a reliable flexible backup.

In this scenario, using a large set of dedicated edge computing devices (each covering a small set of tasks in a limited geographical area) can be superior to using a small set of flexible edge computing devices (each covering more varied tasks over a larger geographical area). As the variety of tasks or the range of geographical areas increase, contentious demands from different types of tasks or regions may knock out an edge device by saturating a particular type of computing resource (e.g., server CPU, low-power CPU, or GPU), causing either hard failure (e.g., a crash or a hang) or a soft failure (e.g., performance-related issues) (Bagchi et al. 2020). Such failures create unnecessary latencies for all covered tasks and regions by requiring the network to fall back on more expensive cloud resources. In contrast, multiple dedicated edge computing devices allow failures to be isolated to limited tasks or geographical regions, therefore they reduce expected usage of more expensive cloud backup. As a result, increasing the number of dedicated edge computing devices can improve network performance as compared with embedding flexibility in existing edge computing devices.

The above network design concept is widely implemented in different forms. For example, GE's Digital Twin allows a company to use less expensive, localized control systems dedicated to different types of turbines, while also allowing the company to use its cloud computing resource to optimize control for all types of turbines (GE Digital 2019). Similarly, the Tsunami Warning Center (TWC) uses local and global seismographic networks to transmit seismograms (Kong et al. 2017). Because latency is a critical concern, localized edge computing devices are installed for each specific susceptible area (rather than using flexible edge devices to cover a wider range of areas, which could cause a larger region to be affected by device failure). In a similar vein, wireless operators using localized devices under mobile edge computing can cheaply deliver content and applications to end users, which provides a better performance than using non-localized edge computing devices that cover a wider area (Brown 2016). These localized devices allow faster and cheaper caching of contents for each target locale, without being affected by failures in neighboring locale. In contrast, flexible edge devices can cover larger service areas, but failures with flexible devices can knock out multiple locales simultaneously, resulting in more expensive backup requests.

6. Conclusion

We study flexibility configuration choices of a supply network that consists of both risky primary suppliers and reliable backup suppliers. We find that adding partial flexibility to this type of supply network is always advantageous. However, the flexibility configuration can have different implications for supply network performance. If flexibility is already installed to primary suppliers, then it is always desirable to add flexibility to backup suppliers as well. In contrast, if flexibility is already available to backup suppliers, then it may not be beneficial to add flexibility to primary suppliers. In fact, we show that a partially flexible network can dominate a fully flexible network, even if flexibility is costless.

We provide a proof that generalizes the two-product result studied by Tomlin and Wang (2005). We show that the irreducibility of partial supply failure states is a necessary and sufficient condition to create opportunities for the supplier diversification benefit to arise in a partially flexible network. Such opportunities, however, do not automatically translate into the supplier diversification benefit – there must exist mechanisms to leverage such opportunities. We show that the presence of backup suppliers does not provide such a mechanism; indeed, without leverage mechanisms, the supplier diversification benefit is never realized for any level of backup capacity, despite the opportunities. Instead, the leverage mechanism must be activated by the flexibility added to backup suppliers. The flexible backup capacity can exploit the opportunities created by the partial failure states to allocate excess capacity to less profitable product demand. Such benefit can be so significant that it can dominate the demand pooling benefit, resulting in a partially flexible network dominating a fully flexible network. This finding refines Tomlin and Wang (2005) who show that, regardless of failure probabilities, a risk-neutral firm always prefers flexible to dedicated suppliers. Our result refines the existing knowledge of flexible resources by pointing out that the availability of backup supply can reverse the findings of Tomlin and Wang (2005), and such a reversal can be quite significant.

From a managerial perspective, we provide intuitive, easy-to-interpret bounds on the supplier diversification and demand pooling benefits. The bounds offer insights and managerial guidance about key drivers of the two benefits associated with a partially flexible network and a fully flexible network. Our analysis suggests that firms need to carefully consider the flexibility configurations of their supply networks because adding flexibility may not be beneficial to firms, and it can even hurt the performance of firms' supply networks. If a firm can gradually reduce the failure risk associated with its primary suppliers, then the firm may find it attractive to reach full flexibility. Otherwise, the firm is better off with a partially flexible network instead of striving for a fully flexible network.

An interesting future research direction is to consider the value of flexibility when supply disruptions can occur at random intervals during a planning horizon. In such a setting, the benefits of flexibility are contingent on when disruptions occur and whether these disruptions overlap among different product lines. Another potential research direction is to explore how flexibility should be configured with more a general, N-product supply network. We hope this research and other future studies help to further our understanding of optimal flexibility configurations in supply networks.

References

Allon, G. and J. A. Van Mieghem (2010). Global dual sourcing: Tailored base-surge allocation to near- and offshore production. *Management Science 56*(1), 110–124.

Ata, B. and J. A. Van Mieghem (2009). The value of partial resource pooling: Should a service network be integrated or product-focused? *Management Science 55*(1), 115–131.

Babich, V. (2010). Independence of capacity ordering and financial subsidies to risky suppliers. *Manufacturing & Service Operations Management 12*(4), 583–607.

Babich, V., A. N. Burnetas, and P. H. Ritchken (2007). Competition and diversification effects in supply chains with supplier default risk. *Manufacturing & Service Operations Management 9*(2), 123–146.

Bagchi, S., M.-B. Siddiqui, P. Wood, and H. Zhang (2020). Dependability in edge computing. *Communications of the ACM 63*(1), 58–66.

Baker, K. R., M. J. Magazine, and H. L. W. Nuttle (1986). The effect of commonality on safety stock in a simple inventory model. *Management Science 32*(8), 982–988.

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

Bassamboo, A., R. S. Randhawa, and J. A. Van Mieghem (2010). Principles on the benefits of manufacturing process flexibility. *Management Science 56*(8), 1285–1303.

Bassamboo, A., R. S. Randhawa, and J. A. Van Mieghem (2012). A little flexibility is all you need: On the asymptotic value of flexible capacity in parallel queuing systems. *Operations Research 60*(6), iii–1565.

Bish, E. K. and Q. Wang (2004). Optimal investment strategies for flexible resources, considering pricing and correlated demands. *Operations Research 52*(6), 954–964.

Boyabatli, O., T. Leng, and L. B. Toktay (2016). The impact of budget constraints on flexible vs. dedicated technology choice. *Management Science 62*(1), 225–244.

Brown, G. (2016). Mobile edge computing use cases & deployment options. *Juniper Networks. https://www.juniper.net/assets/us/en/local/pdf/whitepapers/2000642-en.pdf*.

Chakraborty, T., S. S. Chauhan, and M. Ouhimmou (2016). Mitigating supply disruption with backup supplier under uncertain demand: competition and cooperation. *Working paper CIRRELT-2016-18*.

Chod, J., M. G. Markakis, and N. Trichakis (2021). On the learning benefits of resource flexibility. *Management Science. Published online in Articles in Advance 01 Feb 2021 . https://doi.org/10.1287/mnsc.2020.3795*.

Chopra, S. and M. Sodhi (2004). Managing risk to avoid supply chain breakdown. *MIT Sloan Management Review 46*, 53–61.

DeCroix, G. A. (2013). Inventory management for an assembly system subject to supply disruptions. *Management Science 59*(9), 2079–2092.

Eppen, G. D. (1979). Effects of centralization on expected costs in a multi-location newsboy problem. *Management Science 25*(5), 498–501.

Eppen, G. D., R. K. Martin, and L. Schrage (1989). A scenario approach to capacity planning. *Operations Research 37*(4), 517–527.

Federgruen, A. and N. Yang (2009a). Competition under generalized attraction models: Applications to quality competition under yield uncertainty. *Management Science 55*(12), 2028–2043.

Federgruen, A. and N. Yang (2009b). Optimal supply diversification under general supply risks. *Operations Research 57*(6), 1451–1468.

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

GE Digital (2019). The digital twin: Compressing time to value for digital industrial companies. *https://www.ge.com/digital/lp/digital-twin-compressing-time-value-digital-industrial-companies.*.

Gurler, U. and M. Parlar (1997). An inventory problem with two randomly available suppliers. *Operations Research 45*(6), 904–918.

Harrison, M. J. and J. A. Van Mieghem (1999). Multi-resource investment strategies: Operational hedging under demand uncertainty. *European Journal of Operational Research 113*(1), 17–29.

He, L., Y. Rong, and Z.-J. M. Shen (2019). Product sourcing and distribution strategies under supply disruption and recall risks. *Production and Operations Management. https://doi.org/10.1111/poms.13065*.

Huchzermeier, A. and M. A. Cohen (2006). Valuing operational flexibility under exchange rate risk. *Operations Research 44*(1), 100–113.

Iancu, D. A., E. Ang, and R. Swinney (2017). Disruption risk and optimal sourcing in multi-tier supply networks. *Management Science 63*(8), 2397–2419.

Iravani, S. M. R., B. Kolfal, and M. P. Van Oyen (2007). Call-center labor cross-training: It's a small world after all. *Management Science 53*(7), 1102–1112.

Jayaswala, S., E. Jewkesb, and S. Ray (2011). Product differentiation and operations strategy in a capacitated environment. *European Journal of Operational Research 210*(3), 716–728.

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

Kim, C.-R. (2011). Toyota aims for quake-proof supply chain. *http://www.reuters.com/article/us-toyota-idUSTRE7852RF20110906*.

Kong, L., C. McCreery, and M. Yamamoto (2017). Indian ocean tsunami information center. *http://iotic.ioc-unesco.org/technical- info/tsunami-warning/46/general-information*.

Kouvelis, P. and J. Li (2008). Flexible backup supply and the management of lead-time uncertainty. *Production and Operations Management 17*(2), 184–199.

Netessine, S., G. Dobson, and R. A. Shumsky (2002). Flexible service capacity: Optimal investment and the impact of demand correlation. *Operations Research 50*(2), 375–388.

Pang, G. and W. Whitt (2009). Service interruptions in large-scale service systems. *Management Science 55*(9), 1499–1512.

Parlar, M. and D. Perry (1996). Inventory models of future supply uncertainty with single and multiple suppliers. *Naval Research Logistics 43*, 191–210.

Qi, A., H.-S. Ahn, and A. Sinha (2015). Investing in a shared supplier in a competitive market: Stochastic capacity case. *Production and Operations Management 24*(10), 1537–1551.

Rogers, D. R. and M. White (2013). Key considerations in energy take-or-pay contracts. *King & Spalding Energy Newsletter. https://www.kslaw.com/blog-posts/key-considerations-energy-take-pay-contracts*.

Schmitt, A. J., S. A. Sun, L. V. Snyder, and Z.-J. M. Shen (2015). Centralization versus decentralization: Risk pooling, risk diversification, and supply uncertainty in a one-warehouse multiple-retailer system. *Omega 52*, 201–212.

Sheffi, Y. (2005). *The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage*. Boston: The MIT Press.

Simchi-Levi, D., W. Schmidt, Y. Wei, P. Y. Zhang, K. Combs, Y. Ge, O. Gusikhin, M. Sanders, and D. Zhang (2015). Identifying risks and mitigating disruptions in the automotive supply chain. *Interfaces 45*(5), 375–390.

Snyder, L. V., Z. Atan, P. Peng, Y. Rong, A. J. Schmitt, and B. Sinsoysal (2016). OR/MS models for supply chain disruptions: A review. *IIE Transactions 48*(2), 89–109.

Spinler, S. and A. Huchzermeier (2006). The valuation of options on capacity with cost and demand uncertainty. *European Journal of Operational Research 171*(2), 915–934.

Sting, F. J. and A. Huchzermeier (2010). Ensuring responsive capacity: How to contract with backup suppliers. *European Journal of Operational Research 207*(2), 725–735.

Sun, A. and G. Goldbach (2011). How a flexible supply chain delivers value. *Industry Week. https://www.industryweek.com/supply-chain/procurement/article/21959941/how-a-flexible-supply-chain-delivers-value*.

Thonemann, U. W. and M. L. Brandeau (2000). Optimal commonality in component design. *Management Science 48*(1), 1–19.

Tomlin, B. (2006). On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science 52*(5), 639–657.

Tomlin, B. and Y. Wang (2005). On the value of mix flexibility and dual sourcing in unreliable newsvendor networks. *Manufacturing & Service Operations Management 7*(1), 37–57.

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

Van Mieghem, J. A. (2004). Note: Commonality strategies: Value drivers and equivalence with flexible capacity and inventory substitution. *Management Science 50*(3), 419–424.

Van Mieghem, J. A. (2007). Risk mitigation in newsvendor networks: Resource diversification, flexibility, sharing, and hedging. *Management Science 53*(8), 1269–1288.

Van Mieghem, J. A. and G. Allon (2015). *Operations Strategy* (2nd ed.). Dynamic Ideas LLC.

Wu, S. D., M. Erkoc, and S. Karabuk (2005). Managing capacity in the high-tech industry: A review of literature. *The Engineering Economist 50*(2), 125–158.

Yang, Z., G. Aydin, V. Babich, and D. Beil (2009). Supply disruptions, asymmetric information, and a backup production option. *Management Science 55*(2), 192–209.

Zeng, A. Z. and Y. Xia (2015). Building a mutually beneficial partnership to ensure backup supply. *Omega 52*(2), 77–91.

Online Supplement for "*Product Flexibility Strategy Under Supply and Demand Risk*"

A1. Appendix: Mixed integer programming (MIP) formulations for the *df* and *ff* networks

In this section, we formulate a general scenario based mixed integer programming formulation for the *df* and the *ff* networks. The Bernoulli demand case can then be solved as a special case of the general scenario based formulation.

Let $d_{i,1}$ and $d_{i,2}$, $i=1,\ldots,N$ denote N random draws from the demand distribution for product 1 and 2. Additionally, let M be a large number.

A1.1. MIP formulation for the *df* network

$$
\begin{aligned}
(\mathrm{P_{df}})\ \max_{s,q_1,q_2}\quad & -\gamma s - c(q_1+q_2)(1-\rho) \\
& + \sum_{i=1}^{N} \left[\rho(\pi_1 - v)x_{i,s_1} + (1-\rho)(\pi_1 x_{i,q_1} + (\pi_1 - v)x_{i,\bar{s}_1})\right] \\
& + \sum_{i=1}^{N} \left[(1-\rho)\pi_2 x_{i,q_2} + (\pi_2 - v)(\rho^2 x_{i,s_2} + (1-\rho)^2 x_{i,\bar{s}_2} + \rho(1-\rho)(x_{i,\widehat{s}_2} + x_{i,\widetilde{s}_2})\right] \\
\text{s.t.}\quad & x_{i,s_1} \le s, \quad x_{i,s_1} \le d_{i,1}, \quad i=1,\ldots,N \\
& x_{i,q_1} \le q_1, \quad x_{i,q_1} \le d_{i,1}, \quad i=1,\ldots,N \\
& x_{i,\bar{s}_1} \le s, \quad x_{i,\bar{s}_1} \le z_{i,1}, \quad i=1,\ldots,N \\
& x_{i,q_2} \le q_2, \quad x_{i,q_2} \le d_{i,2}, \quad i=1,\ldots,N \\
& x_{i,s_2} \le d_{i,2}, \quad x_{i,s_2} \le z_{i,2}, \quad i=1,\ldots,N \\
& x_{i,\bar{s}_2} \le z_{i,3}, \quad x_{i,\bar{s}_2} \le z_{i,4}, \quad i=1,\ldots,N \\
& x_{i,\widehat{s}_2} \le d_{i,2}, \quad x_{i,\widehat{s}_2} \le z_{i,3}, \quad i=1,\ldots,N \\
& x_{i,\widetilde{s}_2} \le z_{i,2}, \quad x_{i,\widetilde{s}_2} \le z_{i,4}, \quad i=1,\ldots,N \\
& z_{i,1} \ge d_{i,1} - q_1, \quad z_{i,1} \le (1-b_{i,1})M, \quad z_{i,1} \le d_{i,1} - q_1 + b_{i,1}M, \quad i=1,\ldots,N \\
& z_{i,2} \ge s - d_{i,1}, \quad z_{i,2} \le (1-b_{i,2})M, \quad z_{i,2} \le s - d_{i,1} + b_{i,1}M, \quad i=1,\ldots,N \\
& z_{i,3} \ge s - z_{i,1}, \quad z_{i,3} \le (1-b_{i,3})M, \quad z_{i,3} \le s - z_{i,1} + b_{i,3}M, \quad i=1,\ldots,N \\
& z_{i,4} \ge d_{i,2} - q_2, \quad z_{i,4} \le (1-b_{i,4})M, \quad z_{i,4} \le d_{i,2} - q_2 + b_{i,4}M, \quad i=1,\ldots,N \\
& s, q_1, q_2, x_{i,*}, z_{i,*} \ge 0, \quad i=1,\ldots,N \\
& b_{i,*} \in \{0,1\}, \quad i=1,\ldots,N
\end{aligned}
$$

A1.2. MIP formulation for the ff network

$$
\begin{aligned}
(\mathrm{P}_{\mathrm{ff}}) \max_{s,q} & -fs - cq(1-\rho) \\
& + \sum_{i=1}^{N} [\rho(\pi_1 - v)x_{i,s_1} + (1-\rho)(\pi_1 x_{i,q_1} + (\pi_1 - v)x_{i,\bar{s}_1})] \\
& + \sum_{i=1}^{N} [\rho(\pi_2 - v)x_{i,s_2} + (1-\rho)(\pi_2 x_{i,q_2} + (\pi_2 - v)x_{i,\bar{s}_2})] \\
\text{s.t.} \quad & x_{i,s_1} \le s, \quad x_{i,s1} \le d_{i,1}, \quad i=1,\dots,N \\
& x_{i,q_1} \le q, \quad x_{i,q_1} \le d_{i,1}, \quad i=1,\dots,N \\
& x_{i,\bar{s}_1} \le s, \quad x_{i,\bar{s}_1} \le z_{i,1}, \quad i=1,\dots,N \\
& x_{i,s_2} \le d_{i,2}, \quad x_{i,s_2} \le z_{i,2}, \quad i=1,\dots,N \\
& x_{i,q_2} \le d_{i,2}, \quad x_{i,q_2} \le q - d_{i,1} + z_{i,1}, \quad i=1,\dots,N \\
& x_{i,\bar{s}_2} \le z_{i,3}, \quad x_{i,\bar{s}_2} \le z_{i,4}, \quad i=1,\dots,N \\
& z_{i,1} \ge d_{i,1} - q, \quad z_{i,1} \le (1-b_{i,1})M, \quad z_{i,1} \le d_{i,1} - q + b_{i,1}M, \quad i=1,\dots,N \\
& z_{i,2} \ge s - d_{i,1}, \quad z_{i,2} \le (1-b_{i,2})M, \quad z_{i,2} \le s - d_{i,1} + b_{i,2}M, \quad i=1,\dots,N \\
& z_{i,3} \ge s - z_{i,1}, \quad z_{i,3} \le (1-b_{i,3})M, \quad z_{i,3} \le s - z_{i,1} + b_{i,3}M, \quad i=1,\dots,N \\
& z_{i,4} \ge d_{i,2} - q + d_{i,1} - z_{i,1}, \quad z_{i,4} \le (1-b_{i,4})M,, \quad i=1,\dots,N \\
& z_{i,4} \le d_{i,2} - q + d_{i,1} - z_{i,1} + b_{i,4}M, \quad i=1,\dots,N \\
& s, q, x_{i,*}, z_{i,*} \ge 0, \quad i=1,\dots,N \\
& b_{i,*} \in \{0,1\}, \quad i=1,\dots,N
\end{aligned}
$$

A2. Appendix: Proofs

Proof of Proposition 1. Part (a). The proposition statement can be verified by noting that a flexible backup supplier is able to replicate dedicated backup suppliers by setting the reservation capacity level at $s = s_1 + s_2$. Therefore, a *dd* network can be viewed as a special case of a *df* network, thus *dd* must be weakly dominated by *df*. Part (b). By the proof of Proposition 1, a fully flexible network (ff) dominates a partially flexible primary network (fd) because a flexible backup supplier in ff can replicate any strategies adopted by dedicated backup suppliers in fd. Part (c). By Proposition 1, the partially flexible backup network *df* can perform no worse than the dedicated network *dd*, since any strategies adopted by dedicated backup suppliers can be replicated by a flexible backup supplier. That is, the *df* network can be viewed as a relaxation of the *dd* network, thus *df* must weakly dominate *dd*. □

Proof of Theorem 1. If partial failure states can be reduced to total success or failure states, then the flexible primary supplier can replicate the dedicated primary suppliers by setting $q = q_1 + q_2$.

The flexible supplier, however, does not offer any supplier diversification benefit. Therefore, supplier diversification in a dedicated supply network dd cannot generate any positive value relative to a flexible primary network fd. □

Proof of Proposition 2. Under ff, it is optimal to set $q^{ff} = 2d$ (i.e., no benefit from increasing above $2d$ or decreasing below $2d$). Similarly, under df, it is optimal to order d units of each product from the primary supply (i.e., $q_1^{df} = q_2^{df} = d$). The revenue function (5) under ff at the optimal primary production quantity can be simplified to

$$u_{ff}(s, 2d) = \underbrace{[(\pi_1 - c)\, d + (\pi_2 - c)\, d]\,(1-\rho)}_{(a)} \underbrace{-\gamma s + \left[(\pi_1 - v)\min\{s, d\} + (\pi_2 - v)\,(s-d)^+\right]\rho}_{(b)},$$

where term (a) is related to profit associated with the primary supplier and (b) is the remaining profit associated with the flexible backup supplier. Similarly, the revenue function (4) under df can be simplified to

$$\begin{aligned} u_{df}(s, d, d) =& \underbrace{[(\pi_1 - c)\, d + (\pi_2 - c)\, d]\,(1-\rho)}_{(a)} \underbrace{-\gamma s + \left[(\pi_1 - v)\min\{s, d\} + (\pi_2 - v)\,(s-d)^+\right]\rho}_{(b)} \\ &+ \underbrace{\left[(\pi_2 - v)\left(\min\{s,d\} - (s-d)^+\right)\right]\rho(1-\rho)}_{(c)} \\ =& u_{ff}(s, 2d) + \underbrace{\left[(\pi_2 - v)\left(\min\{s,d\} - (s-d)^+\right)\right]\rho(1-\rho)}_{(c)}, \end{aligned}$$

where terms (a) and (b) have similar interpretations as in the ff network while term (c) is related to profit associated with both the primary and the backup suppliers. Because the optimal backup capacity is $s \le 2d$, it follows that the term (c) is always non-negative, hence the df network dominates the ff network.

Specifically, the optimal reservation quantity for the backup supplier under ff is

$$s^{ff} = \begin{cases} 0, & \rho\,(\pi_1 - v) \le \gamma \\ d, & \rho\,(\pi_2 - v) \le \gamma \le \rho\,(\pi_1 - v) \\ 2d, & \rho\,(\pi_2 - v) \ge \gamma \end{cases}.$$

If $\rho\,(\pi_2 - v) < \gamma < \rho\,(\pi_1 - v)$ (i.e., $s^{ff} = d$), then $u_{df}(s^{ff}, d_1^{df}, d_2^{df}) - u_{ff}(s^{ff}, q^{ff}) = u_{df}(d, d, d) - u_{ff}(d, 2d) = [(\pi_2 - v)\, d]\,\rho\,(1-\rho) > 0$. In this case, df strictly dominates ff by an amount of $(\pi_2 - v)\, d\rho\,(1-\rho)$ whenever $\rho\,(\pi_2 - v) < \gamma < \rho\,(\pi_1 - v)$. In the case when $\gamma \ge \rho(\pi_1 - v)$, $s^{ff} = 0$, hence $u_{df}(s^{ff}, d_1^{df}, d_2^{df}) - u_{ff}(s^{ff}, q^{ff}) = u_{df}(0, d, d) - u_{ff}(0, 2d) = 0$. Therefore, the performance of df and ff are equivalent whenever $\gamma \ge \rho(\pi_1 - v)$. In the case when $\gamma \le \rho(\pi_2 - v)$, $s^{ff} = 2d$, hence $u_{df}(s^{ff}, d_1^{df}, d_2^{df}) - u_{ff}(s^{ff}, q^{ff}) = u_{df}(2d, d, d) - u_{ff}(2d, 2d) = 0$. Consequently, the performance of df and ff are again equivalent whenever $\gamma \le \rho(\pi_2 - v)$. Because the above analysis is based on the optimal solutions to both *df* and ff networks, it follows that the *df* network dominates the ff network when demand is deterministic. □

Proof of Proposition 3. Given that $\tilde{d}_1 = d$ and $\mathsf{E}\left[\tilde{d}_2\right] = d$, the following inequality holds:

$$\mathsf{E}\left[\min\left\{d, \tilde{d}_2 - x\right\}\right] < \mathsf{E}\left[\min\left\{d, \tilde{d}_2\right\}\right] < d \text{ for any } x > 0. \tag{A-1}$$

The support of $\tilde{d}_2$ is $[d_2^-, d_2^+]$. Under deterministic demand for product 1, the optimal production quantity for ff satisfies $q \geq d$. Under this assumption, we have

$$\begin{aligned} u_{ff}(s,q) = & \underbrace{-\gamma s + \mathsf{E}\left[(\pi_1 - v)\min\{d,s\} + (\pi_2 - v)\min\left\{(s-d)^+, \tilde{d}_2\right\}\right]\rho}_{(a)} \\ & + \underbrace{\mathsf{E}\left[-cq + \pi_1 d + \pi_2 \min\left\{\tilde{d}_2, q-d\right\}\right](1-\rho)}_{(b)} \\ & + \underbrace{\mathsf{E}\left[(\pi_2 - v)\min\left\{s, (\tilde{d}_2 - (q-d))^+\right\}\right](1-\rho)}_{(c)}, \end{aligned}$$

where term (a) is related to profit associated with the flexible backup supplier, term (b) is related to profit associated with the primary supplier, and term (c) is related to profit associated with product 2 from both the primary and the backup suppliers.

For *df*, it is optimal to set $q_1 = d$. Then

$$\begin{aligned} u_{df}(s,d,q_2) = & \underbrace{-\gamma s + \mathsf{E}\left[(\pi_1 - v)\min\{d,s\} + (\pi_2 - v)\min\left\{(s-d)^+, \tilde{d}_2\right\}\right]\rho}_{(a)} \\ & + \underbrace{\mathsf{E}\left[-c(d+q_2) + \pi_1 d + \pi_2 \min\left\{\tilde{d}_2, q_2\right\}\right](1-\rho)}_{(b)} \\ & + \underbrace{\mathsf{E}\left[(\pi_2 - v)\begin{pmatrix} \min\left\{s, (\tilde{d}_2 - q_2)^+\right\}(1-\rho)^2 \\ + \min\left\{s, \tilde{d}_2\right\}\rho(1-\rho) \\ + \min\left\{(s-d)^+, (\tilde{d}_2 - q_2)^+\right\}\rho(1-\rho) \\ - \min\left\{(s-d)^+, \tilde{d}_2\right\}\rho(1-\rho) \end{pmatrix}\right]}_{(c')}. \end{aligned}$$

Setting $q_{ff} = q$ and $q_2 = q - d$, then

$$u_{df}(s,d,q-d) - u_{ff}(s,q) = \underbrace{\mathsf{E}\left[(\pi_2 - v)\begin{pmatrix} \min\left\{s, \tilde{d}_2\right\} \\ - \min\left\{s, (\tilde{d}_2 - (q-d))^+\right\} \\ + \min\left\{(s-d)^+, (\tilde{d}_2 - (q-d))^+\right\} \\ - \min\left\{(s-d)^+, \tilde{d}_2\right\} \end{pmatrix}\right]\rho(1-\rho)}_{(c^*)}. \tag{A-2}$$

Notice that

$$\begin{aligned}&\mathsf{E}\left[\min\left\{s,\tilde{d}_2\right\}-\min\left\{s,\left(\tilde{d}_2-\left(q^{ff}(s)-d\right)\right)^+\right\}\right]\\ \geq&\,\mathsf{E}\left[\min\left\{s-d,\tilde{d}_2\right\}-\min\left\{s-d,\left(\tilde{d}_2-\left(q^{ff}(s)-d\right)\right)^+\right\}\right],\end{aligned}$$

where the inequality follows from the fact that the second arguments in each min operator are the same, while the first argument in the LHS is larger than the first argument in the RHS. It then follows directly that $u_{df}(s,d,q-d)-u_{ff}(s,q)\geq 0$.

Next we examine conditions where df strictly dominates ff. For the ff network, the first order condition with respect to q satisfies

$$\frac{\partial u_{ff}(s,q)}{\partial q}=\left[-c+\pi_2 P\left[\tilde{d}_2>q-d\right]-(\pi_2-v)P\left[q-d<\tilde{d}_2\leq s+q-d\right]\right](1-\rho). \tag{A-3}$$

Note that at $q=d$, $\partial u_{ff}(s,d)/\partial q=\left[\pi_2-c-(\pi_2-v)P\left[\tilde{d}_2\leq s\right]\right](1-\rho)$. Therefore,

$$q_{ff}(s)>d \text{ iff } P\left[\tilde{d}_2\leq s\right]<\frac{\pi_2-c}{\pi_2-v}, \tag{A-4}$$

that is, if $c<v$, then $q^{ff}(s)>d$ for any s; and if $c\geq v$, then $q^{ff}(s)>d$ for any $s<s'$ where $s'=G^{-1}\left((\pi_2-c)/(\pi_2-v)\right)$. Given that $v\geq c$, $q^{ff}(s)\geq d$. For $s\geq 0$, $q^{ff}(s)$ satisfies

$$P\left[\tilde{d}_2>q^{ff}(s)-d\right]-\left(\frac{\pi_2-v}{\pi_2}\right)P\left[q^{ff}(s)-d<\tilde{d}_2\leq s+q^{ff}(s)-d\right]=\frac{c}{\pi_2}. \tag{A-5}$$

It follows that $q^{ff\prime}(s)\leq 0$. For df, it is optimal to set $q_1=d$. Let $q_2=q^{ff}(s)-d$ (i.e., $q_1+q_2=q^{ff}(s)$. To compare the ff and df networks for any given backup capacity $0<s<d_2^+$, define univariate function under ff as

$$\begin{aligned}u_{ff}(s)=u_{ff}(s,q^{ff}(s))=&-\gamma s+\mathsf{E}\left[(\pi_1-v)\min\{d,s\}+(\pi_2-v)\min\left\{(s-d)^+,\tilde{d}_2\right\}\right]\rho\\ &+\mathsf{E}\left[\begin{array}{l}-cq^{ff}(s)+\pi_1 d+\pi_2\min\left\{\tilde{d}_2,q^{ff}(s)-d\right\}\\ +(\pi_2-v)\min\left\{s,\left(\tilde{d}_2-(q^{ff}(s)-d)\right)^+\right\}\end{array}\right](1-\rho).\end{aligned} \tag{A-6}$$

Then

$$\begin{aligned}u_{ff}'(s)=&-\gamma+\left((\pi_1-v)\mathbf{I}(s\leq d)+(\pi_2-v)\mathbf{I}(s>d)P\left[\tilde{d}_2>s-d\right]\right)\rho\\ &+\left(\begin{array}{l}q^{ff\prime}(s)\left(\begin{array}{l}-c+\pi_2 P\left[\tilde{d}_2>q^{ff}(s)-d\right]\\ -(\pi_2-v)P\left[q^{ff}(s)-d<\tilde{d}_2\leq s+q^{ff}(s)-d\right]\end{array}\right)\\ +(\pi_2-v)P\left[\tilde{d}_2>s+q^{ff}(s)-d\right]\end{array}\right)(1-\rho).\end{aligned}$$

There are several possible candidate solutions for the optimal s^{ff}, which are enumerated below.

(i) Case of $s^{ff}=d$. Suppose that parameters as such that $s^{ff}=d$, i.e.,

$$u'_{ff}(d)=-\gamma+(\pi_1-v)\rho+\left(\begin{array}{l} q^{ff'}(d)\left(\begin{array}{l}-c+\pi_2 P\left[\tilde{d}_2>q^{ff}(d)-d\right] \\ -(\pi_2-v)P\left[q^{ff}(d)-d<\tilde{d}_2\leq q^{ff}(d)\right]\end{array}\right) \\ +(\pi_2-v)P\left[\tilde{d}_2>q^{ff}(d)\right]\end{array}\right)(1-\rho)=0.$$

Substituting $s^{ff}=d$ into (A-2) yields

$$u_{df}(d,d,q-d)-u_{ff}(d,q)=\mathsf{E}\left[(\pi_2-v)\left(\min\left\{d,\tilde{d}_2\right\}-\min\left\{d,(\tilde{d}_2-q_2)^+\right\}\right)\right]\rho(1-\rho)>0.$$

From (A-1) it follows that the inequality is strict when (A-4) holds (implies $q_2>0$). Thus, *df* strictly dominates ff whenever $s^{ff}=d<G^{-1}\left(\frac{\pi_2-c}{\pi_2-v}\right)$ because $s^{ff}=d<G^{-1}\left(\frac{\pi_2-c}{\pi_2-v}\right)\Rightarrow q^{ff}(d)>d$. Note that if demand of product 2 is deterministic at its expected value $\mathsf{E}\left[\tilde{d}_2\right]=d$, then $q_2=d$ and

$$\begin{aligned} u_{df}(d,d,d)-u_{ff}(d,2d)=&(\pi_2-v)d\rho(1-\rho) \\ >&\mathsf{E}\left[(\pi_2-v)\left(\min\left\{d,\tilde{d}_2\right\}-\min\left\{d,(\tilde{d}_2-q_2)^+\right\}\right)\right]\rho(1-\rho) \\ =&u_{df}(d,d,q-d)-u_{ff}(d,q). \end{aligned}$$

This implies that demand uncertainty reduces the degree of *df* dominance over ff, which proves part (c) of the proposition statement.

(ii) Case of $s\in(0,d)$. The analysis follows analogously from the case of $s=d$ and *df* strictly dominates ff whenever $s^{ff}\in(0,d)$, $s>d_2^- - q_2$, and $s<G^{-1}\left(\frac{\pi_2-c}{\pi_2-v}\right)$.

(iii) Case of $s\in(d,d_2^+]$. It follows directly from (A-2) that $u_{df}(s,d,q-d)-u_{ff}(d,q)\geq 0$. Thus *df* dominates ff whenever $s^{ff}\in(d,d_2^+]$. Note that if $s^{ff}=d_2^+$, then $u_{df}(s,d,q_2)=u_{ff}$.

To sum up, *df* always dominates ff for this case where the demand of product 1 is deterministic and the demand for product 2 is stochastic. This dominance is strict whenever $s^{ff}\in(0,d]$ and $s^{ff}<G^{-1}\left(\frac{\pi_2-c}{\pi_2-v}\right)$. □

Proof of Theorem 2. The necessary part follows from Theorem 1. The sufficient part follows from the fact that irreducible partial failure states can only improve the utilization of the backup capacity, hence the supplier diversification benefit must be non-negative. □

Proof of Proposition 4. If $0<s<\max\left\{\tilde{d}_1+\tilde{d}_2\right\}$, then the primary production satisfies $q_1>0$ and $q_2>0$. Let $s_{1-\rho}$ denote the unit of backup capacity utilized when both primary suppliers succeed and s_ρ denote the unit of backup capacity utilized when both primary suppliers fail. The following relationship then holds

$$s_{1-\rho}<\left\{s_{\rho(1-\rho)},s_{(1-\rho)\rho}\right\}<s_\rho,$$

where $s_{\rho(1-\rho)}$ denotes the state where primary supplier 1 fails, but supplier 2 succeeds, and where $s_{(1-\rho)\rho}$ denotes the state where primary supplier 1 succeeds but supplier 2 fails. The backup capacity allocation in these two partial failure states cannot be decoupled between the two products, because s is flexible. Therefore, these two partial states cannot be absorbed into total success or failure states for each product. □

Proof of Proposition 5. Comparing the terms of the revenue function between df and ff, and looking for remaining terms related with $\rho(1-\rho)$ for the supplier diversification benefit and those related with $(1-\rho)$ for the demand pooling benefit yields the following equality

$$u_{df}(s,q_1,q_2) = u_{ff}(s,q_1+q_2) + \mathrm{SD}(s,q_1,q_2) - \mathrm{DP}(s,q_1,q_2),$$

where

$$\mathrm{DP}(s,q_1,q_2) = \mathsf{E}\left[\begin{array}{l} \quad \pi_1 \min\left\{\tilde{d}_1, q_1+q_2\right\} + \pi_2 \min\left\{\tilde{d}_2, (q_1+q_2-\tilde{d}_1)^+\right\} \\ -\ \pi_1 \min\left\{\tilde{d}_1, q_1\right\} - \pi_2 \min\left\{\tilde{d}_2, q_2\right\} \\ +\ (\pi_1 - v)\min\left\{(\tilde{d}_1 - q_1 - q_2)^+, s\right\} \\ -\ (\pi_1 - v)\min\left\{(\tilde{d}_1 - q_1)^+, s\right\} \\ +\ (\pi_2 - v)\min\left\{\begin{array}{l}\left(s - (\tilde{d}_1 - q_1 - q_2)^+\right)^+, \\ \left(\tilde{d}_2 - (q_1+q_2-\tilde{d}_1)^+\right)^+\end{array}\right\} \\ -\ (\pi_2 - v)\min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, (\tilde{d}_2 - q_2)^+\right\} \end{array}\right](1-\rho)$$

is the demand pooling advantage of ff and

$$\mathrm{SD}(s,q_1,q_2) = (\pi_2 - v)\,\mathsf{E}\left[\begin{array}{l} \quad \min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, \tilde{d}_2\right\} \\ -\ \min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, (\tilde{d}_2 - q_2)^+\right\} \\ +\ \min\left\{(s - \tilde{d}_1)^+, (\tilde{d}_2 - q_2)^+\right\} \\ -\ \min\left\{(s - \tilde{d}_1)^+, \tilde{d}_2\right\} \end{array}\right]\rho(1-\rho)$$

is the supply diversification advantage of df. To interpret the elements of DP, define

$$\begin{array}{ll} x_1^{ff} = \min\left\{\tilde{d}_1, q_1+q_2\right\}, & x_2^{ff} = \min\left\{\tilde{d}_2, (q_1+q_2-\tilde{d}_1)^+\right\}, \\ y_1^{ff} = \min\left\{(\tilde{d}_1 - q_1 - q_2)^+, s\right\}, & y_2^{ff} = \min\left\{\left(s - (\tilde{d}_1 - q_1 - q_2)^+\right)^+, \left(\tilde{d}_2 - (q_1+q_2-\tilde{d}_1)^+\right)^+\right\}, \\ x_1^{df} = \min\left\{\tilde{d}_1, q_1\right\}, & x_2^{df} = \min\left\{\tilde{d}_2, q_2\right\}, \\ y_1^{df} = \min\left\{(\tilde{d}_1 - q_1)^+, s\right\}, & y_2^{df} = \min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, (\tilde{d}_2 - q_2)^+\right\}. \end{array}$$

The above elements correspond to sales volumes given that primary production is operational under both configurations (so that pooling can take place); that is, x_i^{ff} and y_i^{ff} are primary and backup sales, respectively, of product i under ff, while x_i^{df} and y_i^{df} are primary and backup sales, respectively, of product i under df. The following inequalities among volumes are due to the combination of (1) pooling of demand across the total primary production quantity $q = q_1 + q_2$ and (2) prioritizing the allocation of primary production to the most profitable product:

(a) $x_1^{ff} + x_2^{ff} \geq x_1^{df} + x_2^{df}$ (primary sales is larger under ff),

(b) $x_1^{ff} + x_2^{ff} + y_1^{ff} + y_2^{ff} \geq x_1^{df} + x_2^{df} + y_1^{df} + y_2^{df}$ (total sales is larger under ff), and

(c) $x_1^{ff} + y_1^{ff} \geq x_1^{df} + y_1^{df}$ (total sales of the high profit product is higher under ff).

Thus, for any realization of random demands, the following inequality holds:

$$\pi_1 x_1^{ff} + \pi_2 x_2^{ff} + (\pi_1 - v) y_1^{ff} + (\pi_2 - v) y_2^{ff} \geq \pi_1 x_1^{df} + \pi_2 x_2^{df} + (\pi_1 - v) y_1^{df} + (\pi_2 - v) y_2^{df}, \qquad \text{(A-7)}$$

which implies DP ≥ 0. Inequality (A-7) is intuitive due to demand pooling (e.g., no instances of unused primary production of product 1 with simultaneous shortage of primary production of product 2) and the allocation of primary production capacity to the most profitable product (e.g., the constraint requiring that q_i must be allocated to product i demand is relaxed).

To interpret SD, define

$$z_{10} = \min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, \tilde{d}_2\right\}, \quad z_{01} = \min\left\{(s - \tilde{d}_1)^+, (\tilde{d}_2 - q_2)^+\right\},$$
$$z_{11} = \min\left\{\left(s - (\tilde{d}_1 - q_1)^+\right)^+, (\tilde{d}_2 - q_2)^+\right\}, \quad z_{00} = \min\left\{(s - \tilde{d}_1)^+, \tilde{d}_2\right\}.$$

The above elements correspond to backup production volumes of product 2 under df, where the subscripts denote whether the primary suppliers for product 1 or 2 is up (1) or down (0). Note that $z_{10} \geq z_{11}$ (because $\tilde{d}_2 \geq (\tilde{d}_2 - q_2)^+$) and $z_{10} \geq z_{00}$ (because $(\tilde{d}_1 - q_1)^+ \leq \tilde{d}_1$). Further, at least one of the following two identities holds for any parameter values: $z_{01} = z_{11}$ or $z_{01} = z_{00}$. When $(s - \tilde{d}_1)^+ \leq (\tilde{d}_2 - q_2)^+$, then $(s - \tilde{d}_1)^+ \leq \tilde{d}_2$, which implies $z_{01} = z_{00}$. When $(s - \tilde{d}_1)^+ \geq (\tilde{d}_2 - q_2)^+$, then $(s - (\tilde{d}_1 - q_1)^+)^+ \geq (\tilde{d}_2 - q_2)^+$, which implies $z_{01} = z_{11}$. Therefore, $\text{SD}(s, q_1, q_2) \geq 0$. $\square$

Proof of Proposition 6. Part (a). By the corollary assumption, $\tilde{d}_1 = d$ (known demand of product 1). It follows that it is optimal to set $q_1 = d$. Let $q_2 = q - q_1 = q - d$. Then

$$\text{DP}(s, d, q_2) = \mathsf{E}\left[\left(\begin{array}{l} \pi_1 \begin{pmatrix} \min\{d, d + q_2\} \\ -\min\{d, d\} \end{pmatrix} + (\pi_1 - v)\begin{pmatrix} \min\left\{(d - d - q_2)^+, s\right\} \\ -\min\left\{(d-d)^+, s\right\} \end{pmatrix} \\ +\pi_2 \begin{pmatrix} \min\left\{\tilde{d}_2, (d + q_2 - d)^+\right\} \\ -\min\left\{\tilde{d}_2, q_2\right\} \end{pmatrix} \\ +(\pi_2 - v)\begin{pmatrix} \min\left\{\left(s - (d - d - q_2)^+\right)^+, \left(\tilde{d}_2 - (d + q_2 - d)^+\right)^+\right\} \\ -\min\left\{\left(s - (d-d)^+\right)^+, \left(\tilde{d}_2 - q_2\right)^+\right\} \end{pmatrix} \end{array}\right)\right](1 - \rho) = 0.$$

Part (b). If $s=0$, then

$$\mathrm{SD}(0,q_1,q_2)=(\pi_2-v)\,\mathsf{E}\begin{bmatrix}\min\left\{\left(0-\left(\tilde{d}_1-q_1\right)^+\right)^+,\tilde{d}_2\right\}\\+\min\left\{\left(0-\tilde{d}_1\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\\-\min\left\{\left(0-\left(\tilde{d}_1-q_1\right)^+\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\\-\min\left\{\left(0-\tilde{d}_1\right)^+,\tilde{d}_2\right\}\end{bmatrix}\rho(1-\rho)=0.$$

If $s=\max\left\{\tilde{d}_1+\tilde{d}_2\right\}$, then

$$\mathrm{SD}(s,q_1,q_2)=(\pi_2-v)\,\mathsf{E}\begin{bmatrix}\min\left\{s-\left(\tilde{d}_1-q_1\right)^+,\tilde{d}_2\right\}\\+\min\left\{s-\tilde{d}_1,\left(\tilde{d}_2-q_2\right)^+\right\}\\-\min\left\{s-\left(\tilde{d}_1-q_1\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\\-\min\left\{s-\tilde{d}_1,\tilde{d}_2\right\}\end{bmatrix}\rho(1-\rho)=0. \qquad \square$$

Proof of Proposition 7. Define $\mathcal{L}(s)=\mathsf{E}\left[(\pi_1-v)\min\left\{\tilde{d}_1,s\right\}+(\pi_2-v)\min\left\{\left(s-\tilde{d}_1\right)^+,\tilde{d}_2\right\}\right]$, which denotes the expected profit when the primary supply is unavailable. This function depends only on the backup capacity s. The objective function under ff can be re-written as $u_{ff}(s,q)=-\gamma s+\mathcal{L}(s)\rho+v_{ff}(q|s)\,(1-\rho)$, where

$$v_{ff}(q|s)=-cq+\mathsf{E}\begin{bmatrix}\pi_1\min\left\{\tilde{d}_1,q\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q\right)^+,s\right\}\\+\pi_2\min\left\{\tilde{d}_2,\left(q-\tilde{d}_1\right)^+\right\}\\+(\pi_2-v)\min\left\{\left(s-\left(\tilde{d}_1-q\right)^+\right)^+,\left(\tilde{d}_2-\left(q-\tilde{d}_1\right)^+\right)^+\right\}\end{bmatrix}.$$

For any given s, the optimal decision for q can be found by focusing on $v_{ff}(q|s)$. To facilitate profit comparisons, the objective function for df can also be re-written (after some simplification) as $u_{df}(s,q_1,q_2)=-\gamma s+\mathcal{L}(s)\rho+v_{df}(q_1,q_2|s)(1-\rho)$, where

$$v_{df}(q_1,q_2|s)=-c(q_1+q_2)+\mathsf{E}\begin{bmatrix}\pi_1\min\left\{\tilde{d}_1,q_1\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+,s\right\}+\pi_2\min\left\{\tilde{d}_2,q_2\right\}\\+(\pi_2-v)\begin{bmatrix}\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}(1-\rho)\\+\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\tilde{d}_2\right\}\rho\\+\min\left\{\left(s-\tilde{d}_1\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\rho\\-\min\left\{\left(s-\tilde{d}_1\right)^+,\tilde{d}_2\right\}\rho\end{bmatrix}\end{bmatrix}.$$

Similar to the ff case, for any given s, the optimal decision for q_1 and q_2 can be found by focusing on the terms associated with $1-\rho$. If $q_1 = q$, where q is the optimal primary production quantity in ff network, then the profit associated with product 1 is identical between ff and df. Consequently, df dominates ff if there exists q_2^* such that $v_{df}(q, q_2^*|s) > v_{ff}(q|s)$ for any given q. The profit related with product 2 for ff network is

$$v_{2,ff}(q|s) = \pi_2 \,\mathsf{E} \min\left\{\tilde{d}_2, (q-\tilde{d}_1)^+\right\} + (\pi_2 - v)\,\mathsf{E}\min\left\{(s-(\tilde{d}_1-q)^+)^+, (\tilde{d}_2-(q-\tilde{d}_1)^+)^+\right\}.$$

The profit related with product 2 for df network is

$$v_{2,df}(q, q_2|s) = -\,cq_2 + \pi_2\,\mathsf{E}\min\left\{\tilde{d}_2, q_2\right\} + (\pi_2 - v)\,\mathsf{E}\begin{bmatrix} \min\left\{(s-(\tilde{d}_1-q)^+)^+, (\tilde{d}_2-q_2)^+\right\}(1-\rho) \\ +\min\left\{(s-(\tilde{d}_1-q)^+)^+, \tilde{d}_2\right\}\rho \\ +\min\left\{(s-\tilde{d}_1)^+, (\tilde{d}_2-q_2)^+\right\}\rho \\ -\min\left\{(s-\tilde{d}_1)^+, \tilde{d}_2\right\}\rho \end{bmatrix}.$$

Define $\tilde{z} = (q-\tilde{d}_1)^+$, which is the random available quantity for product 2. Then

$$v_{2,ff}(q|s) = \pi_2\,\mathsf{E}\min\left\{\tilde{d}_2, \tilde{z}\right\} + (\pi_2 - v)\,\mathsf{E}\min\left\{\left(s-\left(\tilde{d}_1-q\right)^+\right)^+, \left(\tilde{d}_2-\tilde{z}\right)^+\right\}.$$

Note that $\tilde{z}$ depends only on $\tilde{d}_1$ and q. In addition, by the intermediate value theorem, there exists a unique deterministic z that yields the same profit as $\tilde{z}$ for $v_{2,ff}$. To see this, note that $v_{2,ff}(q|s) = \mathsf{E}\left[g(\tilde{z}\right] \in [\min_{z\in\Omega} g(z), \max_{z\in\Omega} g(z)]$, where $g(\cdot)$ is a continuous function and Ω is the support. Therefore, since $v_{2,ff}(q|s) \in [\min_{z\in\Omega} g(z), \max_{z\in\Omega} g(z)]$ and $g(\cdot)$ is continuous, it follows from the intermediate value theorem that there exists a $z \in \Omega$ such that $g(z) = v_{2,ff}(q|s)$. That is, there exists a unique deterministic z that obtains exactly the same profit as stochastic $\tilde{z}$. Given this observation, let $z = q_2$ denote the deterministic counterpart of $\tilde{z}$, then

$$\mathcal{V}(z) = v_{2,df}(q, q_2|s) - v_{2,ff}(q|s) = -cz + (\pi_2 - v)\rho\,\mathsf{E}\begin{bmatrix} -\min\left\{\left(s-\left(\tilde{d}_1-q\right)^+\right)^+, \left(\tilde{d}_2-z\right)^+\right\} \\ +\min\left\{\left(s-\left(\tilde{d}_1-q\right)^+\right)^+, \tilde{d}_2\right\} \\ +\min\left\{\left(s-\tilde{d}_1\right)^+, \left(\tilde{d}_2-z\right)^+\right\} \\ -\min\left\{\left(s-\tilde{d}_1\right)^+, \tilde{d}_2\right\} \end{bmatrix}$$

Setting $\mathcal{V}(z) \geq 0$ yields the proposition statement. The second part can be proved by showing that $\mathcal{H}(s) \geq 0$. To see this, define two arbitrary non-negative variables A and B, and denote A^- and B^- two corresponding variables that satisfy $A^- \leq A$ and $B^- \leq B$. Then one can verify that

$\min\{A,B\}-\min\{A,B^-\}\geq\min\{A^-,B\}-\min\{A^-,B^-\}$. Applying the above inequality to $\mathcal{H}(s)$ yields the result that $\mathcal{H}(s)\geq 0$. $\square$

Proof of Proposition 8. The proposition statement can be proved in several steps. First, we examine the marginal value of backup capacity s under the ff and df network. We use the same expressions for $\mathcal{L}(s)$ and $u_{ff}(s,q)$ as defined in the proof of Proposition 7. By the envelope theorem, we have

$$\frac{\partial u_{ff}(s,q^*)}{\partial s}=-\gamma+\mathcal{L}'(s)\rho+(1-\rho)\begin{pmatrix}(\pi_1-v)\overline{F}(s+q^*)\\ +\ (\pi_2-v)\int_{\tilde{d}_1\leq s+q^*}\overline{F}_2(s+q^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{pmatrix}, \tag{A-8}$$

where $\mathcal{L}'(s)=(\pi_1-v)\overline{F}_1(s)+(\pi_2-v)\int_{\tilde{d}_1\leq s}\overline{F}_2(s-\tilde{d}_1)dF_1(\tilde{d}_1)$. Similarly, for df network we also use the same definitions of $u_{df}(s,q_1,q_2)$ as developed in the proof of Proposition 7. By the envelope theorem, we have

$$\begin{aligned}\frac{\partial u_{df}(s,q_1^*,q_2^*)}{\partial s}=&-\gamma+\mathcal{L}'(s)\rho^2\\ &+\left[\begin{matrix}(\pi_1-v)\overline{F}_1(s+q_1^*)\\ +\ (\pi_2-v)\begin{pmatrix}\int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s+q_2^*)dF_1(\tilde{d}_1)\\ +\ \int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{pmatrix}\end{matrix}\right](1-\rho)^2\\ &+\left[(\pi_1-v)\overline{F}_1(s+q_1^*)+(\pi_2-v)\begin{pmatrix}\int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s)dF_1(\tilde{d}_1)\\ +\ \int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{pmatrix}\right]\rho(1-\rho)\\ &+\left[(\pi_1-v)\overline{F}_1(s)+(\pi_2-v)\int_{\tilde{d}_1\leq s}\overline{F}_2(s-\tilde{d}_1+q_2^*)dF_1(\tilde{d}_1)\right]\rho(1-\rho).\end{aligned}$$

The above expression can be re-arranged as

$$\begin{aligned}\frac{\partial u_{df}(s,q_1^*,q_2^*)}{\partial s}=&-\gamma+\mathcal{L}'(s)\rho^2+(\pi_1-v)\overline{F}_1(s)\rho(1-\rho)+(\pi_1-v)\overline{F}_1(s+q_1^*)(1-\rho)\\ &+(\pi_2-v)\left[\begin{matrix}\begin{pmatrix}\int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s+q_2^*)dF_1(\tilde{d}_1)\\ +\ \int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{pmatrix}(1-\rho)^2\\ +\begin{pmatrix}\int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s)dF_1(\tilde{d}_1)\\ +\ \int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\\ +\ \int_{\tilde{d}_1\leq s}\overline{F}_2(s-\tilde{d}_1+q_2^*)dF_1(\tilde{d}_1)\end{pmatrix}\rho(1-\rho)\end{matrix}\right].\end{aligned} \tag{A-9}$$

To derive conditions under which df dominates ff, we need to compare $\partial u_{ff}(s,q^*)/\partial s$ with $\partial u_{df}(s,q_1^*,q_2^*)/\partial s$. Instead of making a direct comparison, we first derive the optimal production quantities q^* under the ff network and q_1^* and q_2^* under the df network. We will then use the properties of these optimal production quantities to simplify the comparison of the marginal benefit of backup capacity s on the two network structures.

First consider the ff network. Note that the optimal q is not related to $\mathcal{L}$, thus one can focus on the value function $v_{ff}(q|s)$ (as defined in the proof of Proposition 7). Taking derivatives of $v_{ff}(q|s)$ with respect to q yields

$$\begin{aligned} v'_{ff}(q|s) = & -c+\pi_1\overline{F}_1(q)-(\pi_1-v)\left(F_1(s+q)-F_1(q)\right)+\pi_2\int_{\tilde{d}_1\le q}\overline{F}_2(q-\tilde{d}_1)dF_1(\tilde{d}_1) \\ & +(\pi_2-v)\begin{pmatrix} -\int_{\tilde{d}_1\le q}\left(F_2(s+q-\tilde{d}_1)-F_2(q-\tilde{d}_1)\right)dF_1(\tilde{d}_1) \\ +\int_{\tilde{d}_1>q}^{\tilde{d}_1\le s+q}\overline{F}_2(s+q-\tilde{d}_1)dF_1(\tilde{d}_1) \end{pmatrix}, \end{aligned} \tag{A-10}$$

and $v''_{ff}(q|s) = -(\pi_1-\pi_2)f_1(s+q) - v\int_{\tilde{d}_1\le q} f_2(q-\tilde{d}_1)dF_1(\tilde{d}_1) - (\pi_2-v)\int_{\tilde{d}_1\le s+q} f_2(s+q-\tilde{d}_1)dF_1(\tilde{d}_1)\le 0$. The value function is concave in q and the optimal q satisfies the first order condition, which can be further simplified to $\pi_1-(\pi_1-\pi_2)F_1(s+q)-v\int_{\tilde{d}_1\le q}F_2(q-\tilde{d}_1)dF_1(\tilde{d}_1)-(\pi_2-v)\int_{\tilde{d}_1\le s+q}F_2(s+q-\tilde{d}_1)dF_1(\tilde{d}_1)=c$.

Next consider the *df* network. Similar to the ff network, the optimal q_1 and q_2 are not related to $\mathcal{L}(s)$, and so one can focus on the value function $v_{df}(q_1,q_2|s)$ (see definition in the proof of Proposition 7)). Note that $v_{df}(q_1,q_2|s)$ can be simplified by factoring out the common term $(1-\rho)$ (for $\rho<1$) and combining similar terms,

$$\frac{v_{df}(q_1,q_2|s)}{1-\rho} = -c(q_1+q_2)+\mathsf{E}\left[\begin{array}{l} \quad\pi_1\min\left\{\tilde{d}_1,q_1\right\}+(\pi_1-v)\min\left\{\left(\tilde{d}_1-q_1\right)^+,s\right\} \\ +\ \pi_2\min\left\{\tilde{d}_2,q_2\right\} \\ +\ (\pi_2-v)\left[\begin{array}{l} \quad\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}(1-\rho) \\ +\min\left\{\left(s-\left(\tilde{d}_1-q_1\right)^+\right)^+,\tilde{d}_2\right\}\rho \\ +\min\left\{\left(s-\tilde{d}_1\right)^+,\left(\tilde{d}_2-q_2\right)^+\right\}\rho \end{array}\right] \end{array}\right].$$

Taking derivatives of $v_{df}(q_1,q_2|s)$ with respect to q_1 and q_2 (and ignoring the $(1-\rho)$ term on the left hand side) yields

$$\begin{aligned} \frac{\partial}{\partial q_1}v_{df}(q_1,q_2|s) = & -c+\pi_1\overline{F}_1(q_1)-(\pi_1-v)\left(F_1(s+q_1)-F_1(q_1)\right) \\ & +(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1}\overline{F}_2(s+q_1+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho) \\ & +(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1}\overline{F}_2(s+q_1-\tilde{d}_1)dF_1(\tilde{d}_1)\rho, \end{aligned} \tag{A-11}$$

and

$$\frac{\partial^2}{\partial q_1^2}v_{df}(q_1,q_2|s) = -\pi_1 f_1(q_1)-(\pi_1-v)\left(f_1(s+q_1)-f_1(q_1)\right)$$

$$
\begin{aligned}
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)\\
&+(\pi_2-v)\overline{F}_2(q_2)f_1(s+q_1)(1-\rho)-\overline{F}_2(s+q_2)f_1(q_1)(1-\rho)\\
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1-\tilde{d}_1)dF_1(\tilde{d}_1)\rho\\
&+(\pi_2-v)f_1(s+q_1)\rho-\overline{F}_2(s)f_1(q_1)\rho\\
&\le -vf_1(q_1)-(\pi_1-\pi_2)f_1(s+q_1)\\
&-\overline{F}_2(s+q_2)f_1(q_1)(1-\rho)-\overline{F}_2(s)f_1(q_1)\rho\\
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)\\
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1-\tilde{d}_1)dF_1(\tilde{d}_1)\rho\\
&\le 0.
\end{aligned}
$$

The value function is therefore concave in q_1. Similarly, for q_2 the expression is

$$
\begin{aligned}
\frac{\partial}{\partial q_2}v_{df}(q_1,q_2|s)=&-c+\pi_2\overline{F}_2(q_2)\\
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1}\Big(F_2(s+q_1+q_2-\tilde{d}_1)-F_2(q_2)\Big)\,dF_1(\tilde{d}_1)(1-\rho)\\
&-(\pi_2-v)\int_{\tilde{d}_1\le s}\Big(F_2(s+q_2-\tilde{d}_1)-F_2(q_2)\Big)\,dF_1(\tilde{d}_1)\rho, \qquad \text{(A-12)}
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial^2}{\partial q_2^2}v_{df}(q_1,q_2|s)=&-\pi_2 f_2(q_2)\\
&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1}\Big(f_2(s+q_1+q_2-\tilde{d}_1)-f_2(q_2)\Big)\,dF_1(\tilde{d}_1)(1-\rho)\\
&-(\pi_2-v)\int_{\tilde{d}_1\le s}\Big(f_2(s+q_2-\tilde{d}_1)-f_2(q_2)\Big)\,dF_1(\tilde{d}_1)\rho\\
\le&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)\\
&-(\pi_2-v)\int_{\tilde{d}_1\le s} f_2(s+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)\rho\\
\le&\ 0.
\end{aligned}
$$

The value function is therefore concave in q_2. Further, note that

$$
\begin{aligned}
\frac{\partial^2}{\partial q_1\partial q_2}v_{df}(q_1,q_2|s)=&-(\pi_2-v)\int_{\tilde{d}_1>q_1}^{\tilde{d}_1\le s+q_1} f_2(s+q_1+q_2-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)\\
&+(\pi_2-v)\left(F_2(s+q_2)-F_2(q_2)\right)f_1(q_1)(1-\rho).
\end{aligned}
$$

It can be verified that $\partial^2 v_{df}(q_1,q_2|s)/\partial q_1^2 \times \partial^2 v_{df}(q_1,q_2|s)/\partial q_2^2 \geq (\partial^2 v_{df}(q_1,q_2|s)/\partial q_1 \partial q_2)^2$. The value function is therefore jointly concave in q_1 and q_2. As a result, the first order conditions can be leveraged to derive the structural properties of q^* under the ff network and q_1^* and q_2^* under the df network.

For the ff network, therefore, the optimal q^* satisfies

$$-c+\pi_1\overline{F}_1(q^*)-(\pi_1-v)\left(F_1(s+q^*)-F_1(q^*)\right)+\pi_2\int_{\tilde{d}_1\leq q^*}\overline{F}_2(q^*-\tilde{d}_1)dF_1(\tilde{d}_1)$$
$$+(\pi_2-v)\begin{pmatrix} -\int_{\tilde{d}_1\leq q^*}\left(F_2(s+q^*-\tilde{d}_1)-F_2(q^*-\tilde{d}_1)\right)dF_1(\tilde{d}_1) \\ +\int_{\tilde{d}_1>q^*}^{\tilde{d}_1\leq s+q^*}\overline{F}_2(s+q^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{pmatrix}=0.$$

Substituting the above expression into (A-8) yields

$$\frac{\partial u_{ff}(s,q^*)}{\partial s}=-\gamma+\mathcal{L}'(s)\rho+\Delta_{ff}(1-\rho), \tag{A-13}$$

where $\Delta_{ff}=c-v\left(1-\int_{\tilde{d}_1\leq q^*}F_2(q-\tilde{d}_1)dF_1(\tilde{d}_1)\right)$. Similarly, for the df network, the optimal q_1^* satisfies

$$\begin{aligned}
&-c+\pi_1\overline{F}_1(q_1^*)-(\pi_1-v)\left(F_1(s+q_1^*)-F_1(q_1^*)\right)\\
&+(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)\\
&+(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\rho=0.
\end{aligned}$$

The above expression can be re-arranged as

$$\left\{\begin{array}{l} \left[\begin{array}{l} -c+\pi_1\overline{F}_1(q_1^*)-(\pi_1-v)\left(F_1(s+q_1^*)-F_1(q_1^*)\right) \\ (\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{array}\right](1-\rho)^2 \\ +\left[\begin{array}{l} -c+\pi_1\overline{F}_1(q_1^*)-(\pi_1-v)\left(F_1(s+q_1^*)-F_1(q_1^*)\right) \\ (\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\end{array}\right]\rho(1-\rho)\end{array}\right\}=0$$

$$\Leftrightarrow\left\{\begin{array}{l} \left(-c+\pi_1\overline{F}_1(q_1^*)-(\pi_1-v)\left(F_1(s+q_1^*)-F_1(q_1^*)\right)\right)(1-\rho) \\ +(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)^2 \\ +(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\rho(1-\rho)\end{array}\right\}=0$$

$$\Leftrightarrow\left\{\begin{array}{l} \left(-c+v\overline{F}_1(q_1^*)+(\pi_1-v)\left(\overline{F}_1(s+q_1^*)\right)\right)(1-\rho) \\ +(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)(1-\rho)^2 \\ +(\pi_2-v)\int_{\tilde{d}_1>q_1^*}^{\tilde{d}_1\leq s+q_1^*}\overline{F}_2(s+q_1^*-\tilde{d}_1)dF_1(\tilde{d}_1)\rho(1-\rho)\end{array}\right\}=0$$

Substituting the above expression into (A-9) yields

$$\begin{aligned}
\frac{\partial u_{df}(s,q_1^*,q_2^*)}{\partial s}&=-\gamma+\mathcal{L}'(s)\rho^2+\Delta_{df}(1-\rho)+(\pi_1-v)\overline{F}_1(s)\rho(1-\rho)\\
&+(\pi_2-v)\left[\begin{array}{l} \int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s+q_2^*)dF_1(\tilde{d}_1)(1-\rho)^2 \\ +\int_{\tilde{d}_1\leq q_1^*}\overline{F}_2(s)dF_1(\tilde{d}_1)\rho(1-\rho) \\ +\int_{\tilde{d}_1\leq s}\overline{F}_2(s+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)\rho(1-\rho)\end{array}\right],
\end{aligned} \tag{A-14}$$

where $\Delta_{df} = c - v\overline{F}_1(q_1^*)$. Expanding the $\mathcal{L}'(s)$ term for (A-13) and (A-14) leads to

$$\frac{\partial u_{ff}(s,q^*)}{\partial s} = -\gamma + (\pi_1 - v)\overline{F}_1(s)\rho + \Delta_{ff}(1-\rho) + (\pi_2 - v)\int_{\tilde{d}_1 \le s} \overline{F}_2(s-\tilde{d}_1)dF_1(\tilde{d}_1)\rho, \qquad \text{(A-15)}$$

and

$$\begin{aligned}\frac{\partial u_{df}(s,q_1^*,q_2^*)}{\partial s} =& -\gamma + (\pi_1 - v)\overline{F}_1(s)\rho + \Delta_{df}(1-\rho) \\ &+(\pi_2 - v)\begin{bmatrix} \int_{\tilde{d}_1 \le q_1^*} \overline{F}_2(s+q_2^*)dF_1(\tilde{d}_1)(1-\rho)^2 \\ + \int_{\tilde{d}_1 \le q_1^*} \overline{F}_2(s)dF_1(\tilde{d}_1)\rho(1-\rho) \\ + \int_{\tilde{d}_1 \le s} \overline{F}_2(s+q_2^*-\tilde{d}_1)dF_1(\tilde{d}_1)\rho(1-\rho) \\ + \int_{\tilde{d}_1 \le s} \overline{F}_2(s-\tilde{d}_1)dF_1(\tilde{d}_1)\rho^2 \end{bmatrix}, \end{aligned} \qquad \text{(A-16)}$$

Using the above expressions, the following claim can be established: $\partial u_{df}(s,q_1^*,q_2^*)/\partial s|_{s=0} > \partial u_{ff}(s,q^*)/\partial s|_{s=0}$. To verify this claim, observe that $\partial u_{ff}(s,q^*)/\partial s|_{s=0} = -\gamma + (\pi_1 - v)\rho + \Delta_{ff}(1-\rho)$, and $\partial u_{df}(s,q_1^*,q_2^*)/\partial s|_{s=0} = -\gamma + (\pi_1 - v)\rho + \Delta_{df}(1-\rho) + (\pi_2 - v)\left[F_1(q_1^*)(1-\rho) + F_1(q_1^*)F_2(q_2^*)(1-\rho)^2\right]$. It follows that

$$\begin{aligned}\left.\frac{\partial u_{df}(s,q_1^*,q_2^*)}{\partial s}\right|_{s=0} - \left.\frac{\partial u_{ff}(s,q^*)}{\partial s}\right|_{s=0} &= \left(\Delta_{df} - \Delta_{ff} + (\pi_2 - v)F_1(q_1)(1+F_2(q_2)(1-\rho))\right)(1-\rho) \\ &\ge \left(\Delta_{df} - \Delta_{ff}\right)(1-\rho) \\ &= v\left(F_1(q_1^*) - \int_{\tilde{d}_1 \le q^*} F_2(q^*-\tilde{d}_1)dF_1(\tilde{d}_1)\right)(1-\rho). \end{aligned} \qquad \text{(A-17)}$$

If $q_1^* = q^*$, then the above expression is positive and the claim holds. If $q_1^* < q^*$, then $F_1(q_1^*|s = 0) = 1 - c/\pi_1$, $F_2(q_2^*|s = 0) = 1 - c/\pi_2 < 1 - c/\pi_1$, and $Pr(\tilde{d}_1 + \tilde{d}_2 \le q^*|s = 0 = F_c(q^*|s = 0) \in (1 - c/\pi_2, 1 - c/\pi_1)$. The support of random demand is non-negative, so the convolution F_c can be expressed as $F_c(x) = \int_{x_1=0}^{x} F_2(x - x_1)dF_1(x_1)$. Thus, at $s = 0$ $F_1(q_1^*) - \int_{\tilde{d}_1 \le q^*} F_2(q^* - \tilde{d}_1)dF_1(\tilde{d}_1) = F_1(q_1^*) - F_c(q^*) \ge 0$.

In addition, observe that, all else being equal, $\partial u_{df}(s,q_1^*,q_2^*)/\partial s$ and $\partial u_{ff}(s,q^*)/\partial s$ both are strictly and smoothly decreasing in f and v. Furthermore, $\lim_{s\to\infty}(\partial u_{df}(s,q_1^*,q_2^*)/\partial s - \partial u_{ff}(s,q^*)/\partial s) \to 0$ (see Proposition 6(b)) Thus, there must exist a threshold value of $\gamma + v$, denote as Γ, such that for any $\gamma + v \le \Gamma$, $\partial u_{df}(s,q_1^*,q_2^*)/\partial s \ge \partial u_{ff}(s,q^*)/\partial s$ and that the cumulative attained benefit for the *df* network with backup capacity s^o satisfies

$$\begin{aligned} &\left[u_{df}(s^o,q_1^*,q_2^*) - u_{df}(0,q_1^*,q_2^*)\right] \ge \left[u_{ff}(s^*,q^*) - u_{ff}(0,q^*)\right] + \left[u_{ff}(0,q^*) - u_{df}(0,q_1^*,q_2^*)\right] \qquad \text{(A-18)} \\ \Leftrightarrow\ & u_{df}(s^o,q_1^*,q_2^*) \ge u_{ff}(s^*,q^*). \end{aligned}$$

Claim: There exists a finite Γ such that there exists an s^o such that inequality (A-18) is satisfied. The above claim can be verified by contradiction. Suppose no such s^o exists that satisfies (A-18), so that for any s^o the following inequality holds:

$$\left[u_{df}(s^o,q_1^*,q_2^*) - u_{df}(0,q_1^*,q_2^*)\right] < \left[u_{ff}(s^*,q^*) - u_{ff}(0,q^*)\right] + \left[u_{ff}(0,q^*) - u_{df}(0,q_1^*,q_2^*)\right].$$

However, setting $\Gamma = 0$ as a special case will lead to a sufficiently large s^o such that the above inequality fails to hold. Therefore, there exists a threshold value Γ such that (A-18) at least weakly holds. Moreover, there must exist parameter combinations such that (A-18) strictly holds. To see this, first note that $SD = u_{ff}(0, q^*) - u_{df}(0, q_1^*, q_2^*)$ is independent of s, and bounded above by $\widehat{SD} = u_{ff}(0, q^*) - u_{df}(0, q^*, 0) \leq \pi_2 \mathsf{E} \min\{\tilde{d}_2, (q^* - \tilde{d}_1)^+\}(1-\rho)$. For sufficiently large ρ, $\widehat{SD} = u_{ff}(0, q^*) - u_{df}(0, q^*, 0)$ can be infinitely small, hence a strictly positive Γ will induce a corresponding s^o such that (A-18) strictly holds.

Conversely, when $\gamma + v > \Gamma$, the reverse is true: $u_{df}(s^o, q_1^*, q_2^*) \leq u_{ff}(s^*, q^*)$. This can be seen by the fact that if the above condition fails to hold, then one can simply increment Γ by a small amount until the condition holds. This outcome is assured to happen since as Γ increases, s^o and s^* decrease, and $\lim_{\Gamma \to \infty} s^o, s^* \to 0$ such that $u_{df}(s^o \to 0, q_1^*, q_2^*) \leq u_{ff}(s^* \to 0, q^*)$ is assured to hold.

Next we show that Γ is a decreasing function of c. Observe that

$$\frac{\partial \widehat{SD}}{\partial c} = \frac{\partial u_{ff}(0, q^*) - u_{df}(0, q_1^*, q_2^*)}{\partial c} = (q_1^* + q_2^* - q^*)(1-\rho) \geq 0,$$

where the inequality can be verified by comparing (A-10) with (A-11) and (A-12). It follows that for condition (A-18) to hold, the backup value from s^o for the *df* network must increase and the unit cost of backup capacity $v + f$ must be reduced, which leads to lowered Γ. $\square$

Proof of Proposition 9. The maximum possible value that can be attributed to DP occurs when (a) the primary supplier succeeds (by definition of DP and Proposition 5); (b) the primary production that was originally intended for product 2 is used for product 1 (which can happen under ff but not under *df*); and (c) demand for product 1 exceeds what is “intended” for product 1 had the production been fixed to each product. Combining the above terms together leads to $\overline{\text{DP}}$. The maximum possible value that can be attributed to SD occurs when (a) primary suppliers experience partial failures (by definition of SD and Theorem 2); (b) the backup production is fully utilized to satisfy product 2 demand (with margin $(\pi_2 - v)$ as identified in condition (7) in Proposition 7); and (c) demand for product 1 does not consume backup production. Combining the above terms together leads to $\overline{\text{SD}}$. $\square$