5-2019

Disruption risk mitigation in supply chains: The risk exposure index revisited

Sarah Yini GAO
Singapore Management University, yngao@smu.edu.sg

David SIMCHI-LEVI
Massachusetts Institute of Technology

Chung-Piaw TEO
National University of Singapore

Zhenzhen YAN
Nanyang Technological University

DOI: https://doi.org/10.1287/opre.2018.1776

Disruption Risk Mitigation in Supply Chains - The Risk Exposure Index Revisited

Sarah Yini Gao
Lee Kong Chian School of Business, Singapore Management University, Singapore, yngao@smu.edu.sg

David Simchi-Levi
Institute for Data, Systems, and Society, Department of Civil and Environmental Engineering, and Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, dslevi@mit.edu

Chung-Piaw Teo
Institute of Operations Research and NUS Business School, National University of Singapore, Singapore bizteocp@nus.edu.sg

Zhenzhen Yan
School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, yanzhenzhen8@gmail.com

Simchi-Levi et al. (2014, 2015a) proposed a novel approach using the Time-To-Recover (TTR) parameters to analyze the Risk Exposure Index (REI) of supply chains under disruption. This approach is able to capture the cascading effects of disruptions in the supply chains, albeit in simplified environments – TTRs are deterministic, and at most one node in the supply chain can be disrupted. In this paper, we proposed a new method to integrate probabilistic assessment of disruption risks into the REI approach and measure supply chain resiliency by analyzing the Worst-case CVaR (WCVaR) of total lost sales under disruptions.

We show that the optimal strategic inventory positioning strategy in this model can be fully characterized by a conic program. We identify appropriate cuts that can be added to the formulation to ensure zero duality gap in the conic program. In this way, the optimal primal and dual solutions to the conic program can be used to shed light on comparative statics in the supply chain risk mitigation problem. This information can help supply chain risk managers focus their mitigation efforts on critical suppliers and/or installations that will have a greater impact on the performance of the supply chain when disrupted.

Key words: supply chain risk management; disruption management; time-to-survive; sensitivity analysis; completely positive programming.

---

1. Introduction

" Limited resources mean it is essential to focus risk management efforts where they are most needed and will deliver the biggest benefits."

- Geraint John, Senior Vice President, Research, SCM World, 2014

In 2014, Typhoon Halong hit South East Asia, and wreaked havoc on the supply chains of many companies with operations in this region. The magnitude of the financial impact was not only huge (more than 10 billion, according to 'Apparel," March 12, 2015), but the ripple effects of the disruption reached and affected even companies with no operational footprint in South East Asia. This is not a rare phenomenon confronting supply chain planners; a survey of 151 supply chain executives by Accenture shows that 73% of companies surveyed have experienced supply chain disruptions in the past five years (Ferrer J et al. (2007)). Companies must therefore build more robust and resilient supply chains to cushion their supply chain operations from unforeseen disruptions.

Supply chain disruptions may include events such as fire or machine breakdown in a production facility, an unexpected surge in demand or a reduction in supply, natural disasters, or customs delays in a node of the supply chain. Its impact on performance depends on the system's ability to discover and then recover after the disruption has occurred. Even if we ignore the possibility of long-term damage on facilities and consumer markets, it is nevertheless challenging to model the ripple effects of disruptions as they propagate down the supply chain. Levermann (2014) estimated that a cessation of export in the Philippines due to typhoon Haiyan could affect up to 6% of all US production, reflecting that supply chains have become more interconnected and global.

Hopp and Yin (2006) is an early attempt to analyze how disruption effect propagates in a simplified supply chain (i.e., assembly network), with the additional assumption that only a single node can be disrupted to simplify the model. Simchi-Levi et al. (2014, 2015a) proposed a simpler but novel approach using the Time-To-Recover (TTR) notion to quantify the financial impact of disruptions on the entire supply chain, measured by the Risk Exposure Index (REI). Companies can rank their direct or indirect suppliers using REI to identify the "weak link" in the supply chain. Simchi-Levi et al. (2015b) also introduced the Time-To-Survive (TTS) concept, defined as the maximum length of time the entire supply chain can continue to function normally before the ripple effects of the disruption affect the performance. The notion of TTR, REI and TTS have been implemented in Ford Motor Company, Cisco and the United Nations etc., to manage supply chain risks (cf. Simchi-Levi et al. (2015b)).

In general, there are numerous ways to measure the resiliency of a supply chain. We focus here on the case of lost sales suffered when the supply chain is disrupted. Figure 1 shows the hypothetical performance of a typical supply network during disruption, from onset to final recovery, and the level of sales sustained throughout the disruption. We define a supply chain's TTS to be the initial time interval after disruption during which the supply chain is still capable of serving normal demands; whereas the supply chain's TTR is the time duration between the disruption and the

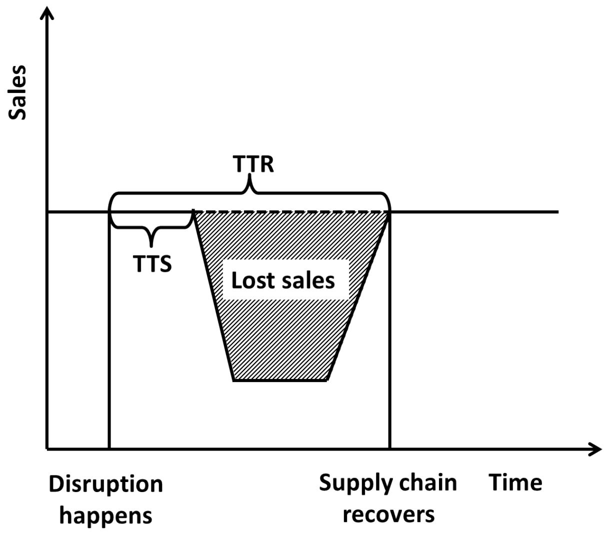


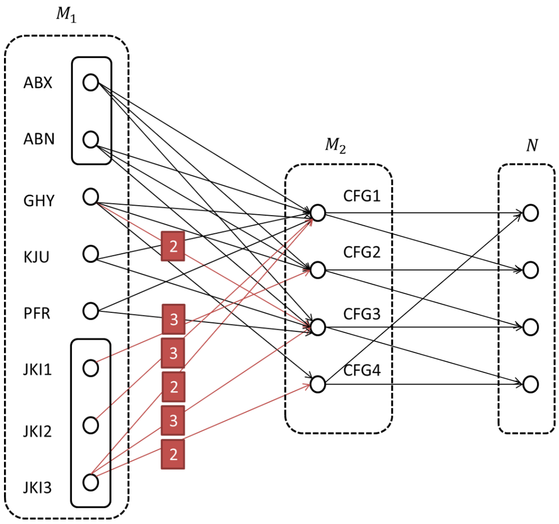
Figure 1 Performance of a supply chain during disruption

time when the supply chain recovers to full functionality. In the rest of this paper, when TTS or TTR is mentioned, it refers to the supply chain's TTS or TTR.

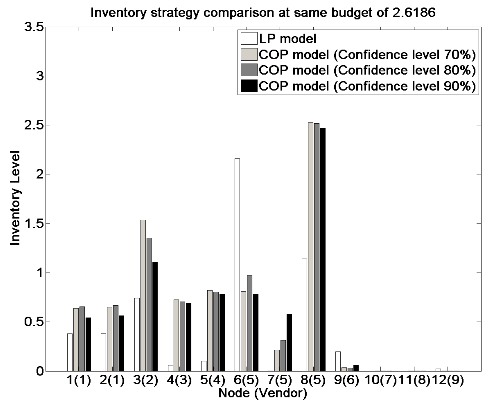
Several measures are identified in Gurnani et al. (2012) as pertinent protection strategies to reduce lost sales, including (i) inventory protection, (ii) capacity protection, (iii) information protection and (iv) supply chain structure design. Note that inventory is one of the most effective risk mitigation strategies used in practice (cf. Geraint, 2014). In this paper, we focus on the use of the inventory protection to cushion the impact of shortages of parts in affected facilities, as these shortages propagate down the supply chain and affect sales to end customers. For instance, in the case of the US West Coast Port lockout in September 2002, manufacturers need to build up inventories in their supply chains in anticipation of the strike, without knowing the duration of the strike and also its impact on other modes of transportation in the network. Our goal in this paper is to develop a model to help guide these inventory protection decisions.

The mitigation strategy adopted is affected by how companies perceive and assess risks in the supply chain. More companies these days are developing early detection capability through weather or social media monitoring, news tracking, and sensor deployment, etc., to sense and respond to supply chain disruptions. CISCO, for instance, uses a six-step incident management system to obtain warning of disruptive events, leading to more accurate and better informed assessment of the disruption risks to its operations (cf. Sheffi (2015)). Tomlin and Snyder (2007) describe how companies like Eaton and UTC have deployed supply-chain monitoring software to "give advance notice of potential supplier instability in time to put safeguards in place." In fact, in August 2004, the system generated a financial alert for a key castings supplier, prompting UTC to increase its inventory buffer as an added layer of protection, averting a disaster heading its way.

As an illustration, consider the following hypothetical supply chain (adapted from Golany (2014)) that will be used throughout our computational study (see Figure 2a): eight key components, ranging from analog display and circuits (ABX and ABN) to different connectors (JK1 to JK3), can be assembled into 4 different configurations (CFG1 to CFG4) and sold in 4 markets (nodes in N). Each configuration in M_2 can be used to serve demands from two markets in N. Items in M_1 are boxed in the figure if they are produced by the same vendor (e.g. ABX and ABN), and hence will be jointly affected if the vendor's facility is disrupted. We call each item node by "plant node" or simply "node"; and each vendor by "vendor node". For instance, ABX is (Plant) Node 1 and it belongs to Vendor Node 1.



(a) The bill-of-material structure is indicated in the graph and the BOM values are shown in the boxes if greater than one.

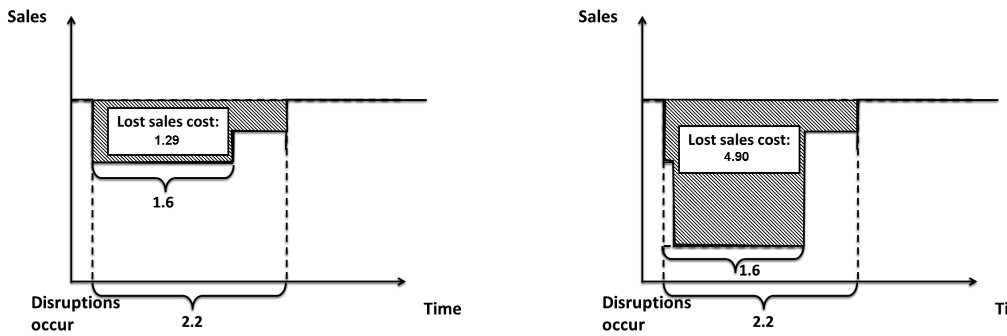


(b) Strategic inventory deployed for each node in the supply chain, based on the LP and COP strategies

Figure 2 Supply Chain Network and the Inventory Deployment using different strategies

What is the optimal way to ramp up inventory in this supply chain when certain disruption risk has been identified? The performance of any mitigation strategy depends essentially on (i) the entire duration of the supply chain disruption, (ii) production loss and time-to-recover at each facility, and (iii) the inventory available in the supply chain to cushion the impact of the disruption. Figure 2b shows the inventory deployment for two classes of mitigation strategies (called "LP" and "COP") studied in this paper. The LP strategy ensures that there will be no lost sales when at most one vendor node in the supply chain is disrupted (as in the REI methodology). This strategy invested a substantial portion of inventory in Node 6 and 8, augmented with a small pool of inventory in Node 9. This is necessary despite the much higher holding cost in vendor node 9, since Node 6 and 8 do not serve Node 9 in the supply chain. With the same budget, the COP strategy opted to invest more in Node 3 and 8, since these are the more flexible nodes in the supply chain, and reduced the inventory investment in Node 9.

The two strategies led to different performance in the risk of lost sales in the supply chain. For instance, when both Vendor Node 4 and 7 are disrupted at the same time (with TTRs of 1.6 and 2.2 respectively), the total lost sales and the resilience curves over time of the two strategies are as shown in Figure 3. Clearly the COP strategy performs better than the LP strategy in this scenario. However, it is foreseeable that the LP strategy will outperform the COP strategy in other disruption scenarios. Is there a formal model to quantify the risk exposure of each inventory deployment strategy?



(a) The resilience curve under COP inventory allocation

(b) The resilience curve under LP inventory allocation

Figure 3 Resilience Curves for the two strategies under disruption

More importantly, for a given inventory mitigation strategy, we need to understand how additional effort can be further directed to "where they are most needed and will deliver the biggest benefits." To do this, we need to understand the comparative statics of the planning parameters on the performance of the supply chain under disruption. For instance, how do we determine whether it is more important to shorten the TTR, increase production rate, or install more inventory at a node? We address these and other issues in the paper. Our main contributions are as follows:

1. **Risk Modeling Using Resilience Curve.** Given the risk assessments, and the accompanying TTR information, we use the framework pioneered by Simchi-Levi et al. (2014, 2015a) to model the cascading effects of disruption on aggregate lost sales in the resilience curve. To account for ambiguity and errors in the risk/probability assessment, we use the "worst-case CVaR" objective to calibrate the performance of the mitigation strategy. More specifically, using a distributionally robust model and assuming only that the first-two moments on disruption probability are known, we show that the optimal inventory allocation under the worst-case CVaR performance of the lost

sales under disruption can be fully characterized by a co-positive program under mild technical conditions. A relaxation of the co-positive cone can be efficiently solved via a semi-definite program.

2. **Disruption Modeling Using Co-positive Cone.** In practice, it is difficult to model the cascading effect of the disruption, especially when these events at different facility may be correlated. We use conic program to model these complicated disruption scenarios, and develop a general approach on adding cuts to ensure that the conic program obtained has no duality gap.

3. **Sensitivity Analysis.** To understand the impact of the key planning parameters on the overall performance of the supply chain under disruption, we show that the optimal solutions of the conic programs can be used to perform rough cut sensitivity analysis on the lost sales sustained during disruption. This exploits a key connection between the conic program and the supply chain disruption problem in the worst-case setting.

4. **Insights on Optimal Protection Strategy.** The optimal mitigation strategy involves a delicate trade-off between the cost and value of inventory at each node. This is in general determined, not only by the operating cost and structure of the supply chain, but also by the budget available to build up these inventories. For instance, for a small budget, the optimal solution may build up inventory on the upstream nodes, increasing such deployments as the budget grows. Surprisingly, beyond certain budget threshold (affected by the TTRs and supply chain structure), the optimal solution will switch to build up inventory at the downstream nodes, and at the same time decrease the inventory deployment at upstream nodes. This feature indicates that supply chain risk mitigation cannot be done purely by looking at the centrality and betweenness of the network structure, as advocated by some scholars (cf. Yan et al. (2015)).

The rest of the paper is organized as follows. In §2, we review the related literature. In §3, we present a general modeling framework and it provides a technique which can be used to reformulate a class of distributionally robust problem into a completely positive program. In §4, we present our risk mitigation model and apply the general framework to determine the optimal inventory allocation across a supply chain with the goal of building a highly resilient supply chain. In §5, by conducting sensitivity analysis on the key parameters, we show how to make use of the solutions obtained in the risk mitigation model to study the effect on lost sales when the supply chain environment changes. In §6, we use a case study to illustrate the performance of our proposed solution. We conclude in §7. All the proofs are presented in the online companion.

2. Literature Review

This paper covers a wide range of topics in supply chain management. We divide our literature review along the key concepts used in our study.

(1) **Inventory Strategy in Risk Mitigation** There are extant literature exploring and presenting different facets of risk management in supply chain. We focus here only on those studies related to the use of inventory as mitigation strategies for supply chain disruption. Meyer et al. (1979) is arguably the first study on supply chain disruption risk mitigation using inventory management strategy, followed by Song and Zipkin (1996), Arreola-Risa and DeCroix (1998), etc. However, most of these papers presented insights only on single product or simplified version of the supply chain. Gurnani et al. (1996) succeeded to extend the problem to two-period, two-component assembly system. More recently, DeCroix (2013) studies the problem under the multiple-period and assembly system settings. Simchi-Levi et al. (2015b), on the other hand, considered a general supply chain network with multiple products. Facing random demands and random disruption events, they adopted a distributionally robust approach to obtain the optimal inventory allocation plan with minimal total inventory from the worst-case perspective. The uncertainty set structure they used is, however, too coarse to capture interdependencies between different nodes in the supply chain. In our paper, we built on the model developed in Simchi-Levi (2015b), but adopt a conic programming approach to solve for the optimal inventory deployment strategy, capturing the correlational structure between different disruption events.

(2) **High-risk Supplier Identification** The assessment of disruption impacts are mostly studied from the perspective of identifying key factors contributing or helping to mitigate the disruption impact (Craighead et al. (2007), Tang (2006), Braunscheidel and Suresh (2009), Kleindorfer and Saad (2005), etc.). Simchi-Levi et al. (2015a), introducing the concepts of TTR, REI and TTS, is one of the first few papers to propose a scheme to rank suppliers by the magnitudes of disruption impacts. Their method is embraced and implemented by several leading businesses. However, this approach assumes only one node is disrupted in each disruption scenario, and hence cannot offer insights when there are breakdowns at multiple nodes in the supply chain. Yan et al. (2015) proposes another method to identify critical suppliers by introducing the concept of a nexus supplier. However, their method is purely based on supply chain network structure and also cannot incorporate the disruption probability information. Our approach is an attempt to fill this gap in the literature, to develop a risk-adjusted approach to find high-risk suppliers based on disruption impacts.

(3) **Completely Positive and Co-positive Programming**. A completely positive program is defined as a linear program over a completely positive cone, whereas a co-positive program is a linear program over a co-positive cone. In general, completely positive and co-positive programs are $\mathcal{NP}$-hard problem. A completely positive cone is defined as

$$\begin{aligned} \mathcal{CP}_n &:= \{M \in S_n | \exists V \in \mathcal{R}_+^{n\times m}, \text{such that } M = VV^T\} \\ &:= \{M \in S_n | \exists \boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_m \in \mathcal{R}_+^n, \text{such that } M = \sum_{i=1}^m \boldsymbol{v}_i \boldsymbol{v}_i^T\} \end{aligned}$$

where, S_n is the set of $n \times n$ symmetric matrices. A co-positive cone is defined as

$$\mathcal{CO}_n := \{M \in S_n | \forall \boldsymbol{v} \in \mathcal{R}^n_+, \boldsymbol{v}^T M \boldsymbol{v} \geq 0\}$$

Completely positive cones and co-positive cones are dual cones to each other. There are rich literature on completely positive and co-positive programming. For more information we refer readers to Berman and Shaked-Monderer (2003). We only present those mostly related to our study. Burer (2009) showed that the well-known $\mathcal{NP}$-hard problem, nonconvex quadratic problems with a mixture of binary and continuous variables has an equivalent completely positive formulation. It is followed by Natarajan et al.(2011), who proposed a *Completely Positive Cross-Moment Model* (CPCMM) giving an equivalent completely positive formulation of the moment-based bound for mixed 0-1 linear programs with random coefficients.

The problem of checking whether a cone is a completely/co-positive cone is already $\mathcal{NP}$ hard. In our paper, we approximate the completely positive cone by a "doubly nonnegative cone (DNN cone)", which is defined to be a positive semi-definite cone with nonnegative entries, i.e. $\{M | M \succcurlyeq 0, M \geq 0\}$. DDN cone is an outer approximation to completely positive cones. For co-positive cones, they are dealt with by an inner approximation, which is the sum of a SDP cone and a nonnegative cone, i.e. $\{M | M = M_1 + M_2, M_1 \succcurlyeq 0, M_2 \geq 0\}$ (cf. de Klerk et al.(2002), Parrilo (2000)).

(4) **Conditional Value-at-Risk (CVaR) and Worst-case CVaR (WCVaR)**. In the case of disruption mitigation, we need to measure the downside risk of lost sales. The well known Value-at-risk (VaR) concept takes such asymmetry into account, but fails to satisfy some natural consistency properties. Fortunately, a related risk measure known as CVaR is shown to be a "coherent risk measure" (Pflug (2000)). Furthermore, Rockafellar and Uryasev (2000, 2002) showed that CVaR can be equivalently solved as a convex optimization problem. In this paper, we use CVaR to quantify the disruption risk on lost sales. Moreover, in order to account for the ambiguity and errors in the risk/probability assessment, we apply distrubitionally robust approach and study this risk measure under the worst-case distribution given partial moment information.

Specifically, CVaR and VaR can be defined as follow. Let $Z(\boldsymbol{r}, \tilde{\boldsymbol{v}})$ denote the lost sales associated with decision $\boldsymbol{r}$ and random disruption event $\tilde{\boldsymbol{v}}$. Given a confidence level $1 - \eta$, say $\eta = 0.05$ (i.e. at 95% confidence level), let

$$VaR_{1-\eta}(\boldsymbol{r}) := \text{argmin}\left\{ t \,\middle|\, P\Big(Z(\boldsymbol{r}, \tilde{\boldsymbol{v}}) \geq t \Big) \leq \eta \right\}.$$

$CVaR_{1-\eta}(\boldsymbol{r})$ is defined as

$$\frac{1}{\eta}\mathbf{E}\left[Z(\boldsymbol{r}, \tilde{\boldsymbol{v}}) \,\middle|\, Z(\boldsymbol{r}, \tilde{\boldsymbol{v}}) \geq VaR_{1-\eta}(\boldsymbol{r}) \right].$$

Rockafellar and Uryasev (2000, 2002) show that CVaR can be equivalently solved by minimizing an auxiliary convex function with respect to the variable θ, i.e.,

$$CVaR_{1-\eta}(\boldsymbol{r}) := \min_{\theta} \theta + \frac{1}{\eta}\mathbf{E}\left[(Z\left(\boldsymbol{r}, \tilde{\boldsymbol{v}}) - \theta\right)^{+}\right].$$

The work by Rockafellar and Uryasev (2000, 2002) has helped to popularize the use of CVaR as a replacement for the computationally intractable notion of VaR. Incorporating the flavor of distributionally robust optimization, $WCVaR_{1-\eta}(\boldsymbol{r})$ is defined as

$$\max_{p(\tilde{v})\in\mathcal{P}} \left\{\min_{\theta} \theta + \frac{1}{\eta}\mathbf{E}\left[(Z(\boldsymbol{r}, \tilde{\boldsymbol{v}}) - \theta)^{+}\right]\right\}.$$

where $\mathcal{P}$ is a family of distributions that can be used to describe the risk profile $\tilde{\boldsymbol{v}}$. Applying the minmax theorem, Zhu and Fukushima (2009) further showed that it can be equivalent reformulated as

$$\min_{\theta} \left\{\theta + \frac{1}{\eta}\max_{p(\tilde{v})\in\mathcal{P}} \mathbf{E}\left[Z(\boldsymbol{r}, \tilde{\boldsymbol{v}}) - \theta)^{+}\right]\right\}.$$

In our paper, we will adopt the above WCVaR formulation in dealing with lost sales due to disruption, using only the means and covariances of node disruption to capture the interdependencies between disruption events.

3. General Completely Positive and Co-positive Framework

Before we introduce the supply chain risk mitigation model, we firstly present the main modelling methodology we apply in the paper. This general modeling framework provides a technique which can be used to reformulate a class of distributionally robust problem into a completely positive program. It will be clear to the reader later that the supply chain risk mitigation model is an application of this framework.

Specifically, consider a stochastic mixed 0-1 quadratic problem with complimentary constraint as follows.

$$\begin{array}{lll} Z(\tilde{\boldsymbol{v}}) = \max & \boldsymbol{c}_1^{\top}\boldsymbol{x} + \tilde{\boldsymbol{v}}^{\top}C_2\boldsymbol{x} + \boldsymbol{x}^{\top}C_3\boldsymbol{x} & \\ s.t. & A_1\boldsymbol{x} = \boldsymbol{b}_1 & \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2 - M\tilde{\boldsymbol{v}} & \\ & (A_3\boldsymbol{x}) \circ (A_4\boldsymbol{x}) = 0 & \\ & x_j \in \{0,1\} & \forall j \in \mathcal{B} \\ & \boldsymbol{x} \geq 0 & \end{array} \tag{1}$$

Here, both the linear objective coefficients and the right-hand side parameters are random, modelled using a random vector, $\tilde{\boldsymbol{v}}$. Note that $\boldsymbol{x} \in \mathbb{R}_{+}^{n}$ is the decision variables; and $\mathcal{B}$ is the set of indices of binary decision variables. All other parameters such as $\boldsymbol{c}_j$, $\boldsymbol{b}_j$, C_j, A_j and M are pre-determined inputs. $\circ$ indicates the elementary multiplication.

Given the distribution of random vector $\tilde{\boldsymbol{v}}$, it is of interest to find the expected value of the optimal objective $\mathbf{E}[Z(\tilde{\boldsymbol{v}})]$. In our paper, we adopt the distributional robustness concept, and assume only the first-two moments and support information of randomness are known. Specifically, assume $\tilde{\boldsymbol{v}}$ lies in the support set $\mathcal{D}$, with $\boldsymbol{\mu} = \mathbf{E}[\tilde{\boldsymbol{v}}]$, $\Sigma = \mathbf{E}[\tilde{\boldsymbol{v}}\tilde{\boldsymbol{v}}^{\mathsf{T}}]$. The central problem we solve is

$$Z^m = \max_{\tilde{\boldsymbol{v}} \sim (\mathcal{D}, \boldsymbol{\mu}, \Sigma)} \mathbf{E}\left[Z(\tilde{\boldsymbol{v}})\right] \tag{2}$$

We further make the following assumptions on the inner maximization problem.

(A1) The feasible region is not empty and is bounded;

(A2) If decision variables satisfy the linear constraints, then those decision variables with indices in $\mathcal{B}$ are between 0 and 1 and they also satisfy $A_3\boldsymbol{x} \geq 0$ and $A_4\boldsymbol{x} \geq 0$.

This class of distributionally robust problem is general enough to include some of the problems studied in literature as special cases (cf. Burer (2009) and Natarajan et al.(2011)). We will show in this section that under assumptions (A1) to (A2), when the uncertainty set has certain special structures that can be modelled using complementary and linear equations, Problem (2) can also be equivalently reformulated as a completely positive program.

To reformulate the uncertainty set as a completely positive program, we need to address the feasibility issue of the presumed moments.

Definition 1. (Bertsimas and Sethuraman (2000)) A sequence $(\boldsymbol{\mu}, \Sigma)$ is a feasible $(n, 2, \mathcal{D})$-moment sequence if there is a multivariate random variable $\boldsymbol{v} = (v_1, ..., v_n)$ with domain $\mathcal{D} \subseteq R^n$, whose moments are given by $(\boldsymbol{\mu}, \Sigma)$, that is $\mu_k = \mathbf{E}[v_k], \Sigma_{ij} = \mathbf{E}[v_i v_j], \forall k, i, j = 1, ..., n$.

The theory of moments attempts to characterize valid moment sequences using semidefinite program. Specifically, when $\mathcal{D} = R^n$, the first-two valid moment sequences can be exactly represented by a semidefinite matrix. Natarajan et al.(2001)'s result implies when $\mathcal{D} = R^n_+$, the first-two valid moments can be equivalently characterized by a completely positive matrix. The following proposition gives the conditions on the support set $\mathcal{D}$ and the conditions on feasible moment sequences such that the uncertainty set can be represented by a completely positive matrix.

Proposition 1. *When the support set $\mathcal{D}$ takes the following form,*

$$\mathcal{D} = \left\{ \tilde{\boldsymbol{v}} \in \mathcal{R}^n_+ \;\middle|\; \begin{array}{l} M_1\tilde{\boldsymbol{v}} = \boldsymbol{b} \\ (M_2\tilde{\boldsymbol{v}}) \circ (M_3\tilde{\boldsymbol{v}}) = 0 \\ \tilde{v}_i \in \{0,1\}, \forall i \in \mathcal{U}^{\mathcal{B}} \end{array} \right\}$$

and under the following assumptions,

(AM1) $\mathcal{D}$ is nonempty and bounded (Endnote 1.)

(AM2) For all $\boldsymbol{v} \geq 0$ satisfying $M_1\boldsymbol{v} = \boldsymbol{b}$, we have $v_i \in [0,1]$ for all $i \in \mathcal{U}^{\mathcal{B}}$, $M_2\boldsymbol{v} \geq 0$ and $M_3\boldsymbol{v} \geq 0$, the sequence $(\boldsymbol{\mu}, \Sigma)$ is a feasible $(n, 2, \mathcal{D})$-moment sequence if and only if the following is not empty.

$$\mathcal{M} = \left\{ (\boldsymbol{w}, X) \;\middle|\; \begin{array}{ll} \boldsymbol{w} = \boldsymbol{\mu}, & X = \Sigma, \\ M_1\boldsymbol{w} = \boldsymbol{b}, & diag(M_1 X M_1^{\mathsf{T}}) = \boldsymbol{b} \circ \boldsymbol{b}, \\ w_i = X_{ii}, \forall i \in \mathcal{U}^{\mathcal{B}} & diag(M_2 X M_3^{\mathsf{T}}) = 0, \\ \begin{pmatrix} 1 & \boldsymbol{w}^{\mathsf{T}} \\ \boldsymbol{w} & X \end{pmatrix} \succcurlyeq_{cp} 0 & \end{array} \right\}$$

where diag($\cdot$) denotes the diagonal elements of a matrix.

An example of such an uncertainty set is a family of multivariate Bernoulli distributions with given first-two moments. In this case, the support set is as follows.

$$\mathcal{D}(\mathcal{B}) = \left\{ \tilde{\boldsymbol{v}} \in \mathbb{R}_+^n \;\middle|\; \tilde{v}_i \in \{0,1\}, \forall i = 1,...,n \right\}$$

Consequently, we would have the following characterization of the feasible moment sequence $(\boldsymbol{\mu}, \Sigma)$.

$$\mathcal{M}(\mathcal{B}) = \left\{ (\boldsymbol{w}, X) \;\middle|\; \begin{array}{ll} \boldsymbol{w} = \boldsymbol{\mu}; & X = \Sigma; \\ w_i = X_{ii}, \forall i = 1,...,n; & s_i = X^s_{ii}, \forall i = 1,...,n; \\ w_i + s_i = 1, \forall i = 1,...,n; & X_{ii} + X^s_{ii} + 2Y^s_{ii} = 1, \forall i = 1,...,n; \\ \begin{pmatrix} 1 & \boldsymbol{w}^\intercal & \boldsymbol{s}^\intercal \\ \boldsymbol{w} & X & Y^s \\ \boldsymbol{s} & {Y^s}^\intercal & X^s \end{pmatrix} \succcurlyeq_{cp} 0 & \end{array} \right\}$$

In the following, we will show Problem (2) has an equivalent completely positive program formulation under the assumptions (A1) to (A2) and when the uncertainty set satisfies the requirements in Proposition (1) Note that our result can be generalized to the case that only a subvector of random variable is specified with moment information. Consider an a subvector of $\tilde{\boldsymbol{v}}$ with indices in $\mathcal{U}$ with $\mathbf{E}[\tilde{\mathbf{v}}_\mathbf{i}] = \mu_\mathbf{i}$ and $\mathbf{E}[\tilde{\mathbf{v}}_\mathbf{i}\mathbf{v}_\mathbf{j}] = \boldsymbol{\Sigma}_\mathbf{ij}$ for all $i, j \in \mathcal{U}$. To simplify the notation and discussion, we focus on the following problem.

$$Z^m = \max_{\tilde{\boldsymbol{v}} \sim (\mathcal{D}, \mu_i, \Sigma_{ij}, \forall i,j \in \mathcal{U}} \mathbf{E}[Z(\tilde{\boldsymbol{v}})] \quad \text{when } \mathcal{D} = \left\{ \tilde{\boldsymbol{v}} \in \mathcal{R}_+^n \;\middle|\; \begin{array}{l} M_1 \tilde{\boldsymbol{v}} = \boldsymbol{b} \\ \tilde{v}_i \in \{0,1\} \ \forall i \in \mathcal{U}^\mathcal{B} \end{array} \right\} \tag{3}$$

Let $\boldsymbol{x(v)}$ be the optimal solutions to the inner maximization problem of Problem (3) for a realization $\boldsymbol{v}$. It is possible that there are multiple optimal solutions in the support of strictly positive measure. We define $\boldsymbol{x(v)}$ to be a randomly selected optimal solution at $\boldsymbol{v}$, and

$$\begin{gathered} \boldsymbol{p^x} := \mathbf{E}[\boldsymbol{x}(\tilde{\boldsymbol{v}})], \quad \boldsymbol{p^w} := \mathbf{E}[(\tilde{\boldsymbol{v}})], \quad \boldsymbol{p^s} := \mathbf{E}[(\tilde{\boldsymbol{s}})] \\ X^x := \mathbf{E}[\boldsymbol{x}(\tilde{\boldsymbol{v}})\boldsymbol{x}(\tilde{\boldsymbol{v}})^T], \quad X^w := \mathbf{E}[\tilde{\boldsymbol{v}}\tilde{\boldsymbol{v}}^T], \quad X^{ws} := \mathbf{E}[\tilde{\boldsymbol{v}}\tilde{\boldsymbol{s}}^T] \\ Y^x := \mathbf{E}[\boldsymbol{x}(\tilde{\boldsymbol{v}})\tilde{\boldsymbol{v}}^T], \; Y^{xs} := \mathbf{E}[\boldsymbol{x}(\tilde{\boldsymbol{v}})\tilde{\boldsymbol{s}}^T], \; Y^{ws} := \mathbf{E}[\tilde{\boldsymbol{v}}\tilde{\boldsymbol{s}}^T] \end{gathered}$$

The objective function in (2) can be written as

$$\mathbf{E}\left[\boldsymbol{c}_1{}^\intercal \boldsymbol{x}(\tilde{\boldsymbol{v}}) + C_2 \cdot \boldsymbol{x}(\tilde{\boldsymbol{v}})\tilde{\boldsymbol{v}}^T + C_3 \cdot \boldsymbol{x}(\tilde{\boldsymbol{v}})\boldsymbol{x}(\tilde{\boldsymbol{v}})^T\right] = \boldsymbol{c}_1{}^\intercal \boldsymbol{p}^x + C_2 \cdot Y^x + C_3 \cdot X^x.$$

where $\cdot$ represents inner product. We reformulate the distributionally robust problem into a completely positive cone problem in Theorem 1.

THEOREM 1. *Problem (3) is equivalent to the following completely positive program.*

$$
\begin{array}{llll}
Z^m = & \max \boldsymbol{c_1}^\intercal \boldsymbol{p}^x + C_2 \cdot Y^x + C_3 \cdot X^x & & \\
& s.t.\ \underline{\textit{Constraints on Decision Variables}} & & \underline{\textit{Dual Variables}} \\
& A_1 \boldsymbol{p}^x = \boldsymbol{b}_1 & & \boldsymbol{\phi}_x \\
& diag(A_1 X^x A_1^\intercal) = \boldsymbol{b}_1 \circ \boldsymbol{b}_1 & & \boldsymbol{\epsilon}_x \\
& A_2 \boldsymbol{p}^x + M\mathbf{w} = \boldsymbol{b}_2 & & \boldsymbol{\phi}_{xv} \\
& diag\Big(\begin{pmatrix} A_2 & M \end{pmatrix} \begin{pmatrix} \boldsymbol{X}^x & Y^x \\ Y^{x\intercal} & X^w \end{pmatrix} \begin{pmatrix} A_2 & M \end{pmatrix}^\intercal \Big) = \boldsymbol{b}_2 \circ \boldsymbol{b}_2 & & \boldsymbol{\epsilon}_{xv} \\
& diag(A_3 X^x A_4^\intercal) = \mathbf{0} & & \boldsymbol{\lambda} \\
& p_j^x = X_{jj}^x, & \forall j \in \mathcal{B} & \boldsymbol{\psi}_x \\
& & & \\
& \underline{\textit{Constraints on Random Variables}} & & \\
& M_1 \boldsymbol{w} = \boldsymbol{b} & & \boldsymbol{\phi}_v \\
& diag(M_1 X^w M_1^\intercal) = \boldsymbol{b} \circ \boldsymbol{b} & & \boldsymbol{\epsilon}_v \\
& w_j = X_{jj}^w, & \forall j \in \mathcal{U}^{\mathcal{B}} & \boldsymbol{\psi}_v \\
& \boldsymbol{w}_j = \mu_j, & \forall j \in \mathcal{U} & \nu_v \\
& \boldsymbol{X}_{ij}^w = \Sigma_{ij}, & \forall i,j \in \mathcal{U} & \Theta_v \\
& & & \\
CP = & \begin{pmatrix} 1 & \boldsymbol{w}^\intercal & \boldsymbol{p}^{x\intercal} \\ \boldsymbol{w} & X^w & Y^{x\intercal} \\ \boldsymbol{p}^x & Y^x & X^x \end{pmatrix} \succcurlyeq_{cp} 0 & & \rho
\end{array}
\tag{4}
$$

where $\mathcal{U}$ *denotes the set of random variables with specified moments. We assume there is a partition of principal sub-matrices of* X^w *specified with moments.* $\mathcal{U}^{\mathcal{B}}$ *denotes the set of Bernoulli random variable.*

The proof applies similar techniques employed in the proof of the main result in Natarajan et al.(2011) and the details are given in the online companion. Note that the constraint $\text{diag}(A_3 X^x A_4^\intercal) = \mathbf{0}$ can also be written as $\text{diag}(A_4 X^x A_3^\intercal) = \mathbf{0}$. This is needed to obtain symmetry in the dual formulation. The corresponding dual co-positive program can be written as follows.

$$
\begin{array}{l}
\min \rho + \mu^\intercal \nu_v + \Sigma \cdot \Theta_v + \boldsymbol{b}_1^\intercal \boldsymbol{\phi}_x + \boldsymbol{b}_1^\intercal (\boldsymbol{\epsilon}_x \circ \boldsymbol{b}_1) + \boldsymbol{b}_2^\intercal \boldsymbol{\phi}_{xv} + \boldsymbol{b}_2^\intercal (\boldsymbol{\epsilon}_{xv} \circ \boldsymbol{b}_2) + \boldsymbol{b}^\intercal \boldsymbol{\phi}_v + \boldsymbol{b}^\intercal (\boldsymbol{\epsilon}_v \circ \boldsymbol{b}) \\
s.t. \begin{pmatrix} \rho & \frac{1}{2}(\nu_v + \boldsymbol{\psi}_v + M^\intercal \boldsymbol{\phi}_{xv} + M_1^\intercal \boldsymbol{\phi}_v)^\intercal & \frac{1}{2}(B^\intercal \boldsymbol{\psi}_x + {A_1}^\intercal \boldsymbol{\phi}_x + A_2^\intercal \boldsymbol{\phi}_{xv} - \boldsymbol{c}_1)^\intercal \\ \frac{1}{2}(\nu_v + \boldsymbol{\psi}_v + M^\intercal \boldsymbol{\phi}_{xv} + M_1^\intercal \boldsymbol{\phi}_v) & \Theta_v + M^\intercal \Lambda(\boldsymbol{\epsilon}_{xv}) M + M_1^\intercal \Lambda(\boldsymbol{\epsilon}_v) M_1 - \Lambda(\boldsymbol{\psi}_v) & (A_2^\intercal \Lambda(\boldsymbol{\epsilon}_{xv}) M - \frac{C_2}{2})^\intercal \\ \frac{1}{2}(B^\intercal \boldsymbol{\psi}_x + {A_1}^\intercal \boldsymbol{\phi}_x + A_2^\intercal \boldsymbol{\phi}_{xv} - \boldsymbol{c}_1) & A_2^\intercal \Lambda(\boldsymbol{\epsilon}_{xv}) M - \frac{C_2}{2} & {A_1}^\intercal \Lambda(\boldsymbol{\epsilon}_x) A_1 + A_2^\intercal \Lambda(\boldsymbol{\epsilon}_{xv}) A_2 + A_4^\intercal \Lambda(\boldsymbol{\lambda}) A_3 + A_3^\intercal \Lambda(\boldsymbol{\lambda}) A_4 - B^\intercal \Lambda(\boldsymbol{\psi}_x) B - C_3 \end{pmatrix} \\
\quad \succcurlyeq_{co} 0
\end{array}
\tag{5}
$$

where $\Lambda(\boldsymbol{u})$ transforms the vector $\boldsymbol{u}$ into a diagonal matrix with all off-diagonal entries equal 0, and B is a diagonal matrix with 1 in the (j, j) entry if the decision variable x_j is binary, 0 otherwise.

3.1. Conic Strong Duality

Strong duality between Problem (4) and Problem (5) does not hold in general. In the following, we present a sufficient condition on the problem structure to ensure strong duality holds. The key idea is to construct an interior point in the co-positive cone in (5). Then the strong duality holds according to Slater condition. Our construction is inspired by Hanasusanto and Kuhn (2017).

Lemma 1 **(Co-positive Schur Complement (Hanasusanto and Kuhn (2017)))**. *Consider a symmetric matrix*

$$D = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}$$

with $A \succ 0$. Then $D \succ_{co} 0$ if $C - B^\top A^{-1} B \succ_{co} 0$.

Proposition 2. *Consider the completely positive program (4) and its dual co-positive program (5). If $(A_1^\top A_1 - C_3) \succ_{co} 0$, then there is no duality gap between the two problem.*

When the random cost function is linear (i.e. $C_3 = \mathbf{0}$) and the constraint matrix A_1 is non-negative with no zero column, the condition holds and the distributionally robust problem studied in this paper has no duality gap using this framework. This includes for instance the maximum order statistic problem as a special case.

In the case when the objective function is quadratic in the decision variables (i.e. $C_3 \neq \mathbf{0}$), we can remove the dependence on the cost efficient C_3 by exploiting the problem structure to derive similar strong duality result.

Proposition 3. *If the decision variable $\boldsymbol{x}$ can be decomposed into two parts $\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}$, such that (i) the objective of Problem (1) can be rewritten as $\hat{\boldsymbol{c}}_1^\top \boldsymbol{x}_1 + \boldsymbol{v}^\top \hat{C}_2 \boldsymbol{x}_1 + \boldsymbol{x}_2^\top \hat{C}_3 \boldsymbol{x}_1$, (ii) the linear constraint $A_1 \boldsymbol{x} = \boldsymbol{b}_1$ can be decomposed to $A_1^1 \boldsymbol{x}_1 = \boldsymbol{b}_1^{(1)}$ and $A_1^2 \boldsymbol{x}_2 = \boldsymbol{b}_1^{(2)}$, and (iii) the coupling constraint $A_2 \boldsymbol{x} + M\boldsymbol{w} = \boldsymbol{b}_2$ can be rewritten as $A_2^2 \boldsymbol{x}_2 + M\boldsymbol{w} = \boldsymbol{b}_2$, then strong duality holds provided $A_1^{1\top} A_1^1 \succ_{co} 0$, and $A_1^{2\top} A_1^2 + A_2^{2\top} A_2^2 \succ 0$.*

To establish the condition $A_1^{1\top} A_1^1 \succ_{co} 0$, the following observation will be useful.

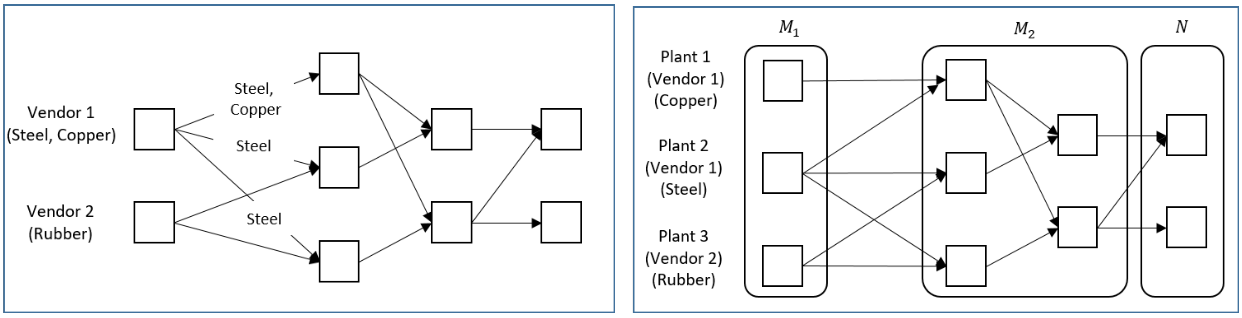
Lemma 2 **(Lemma 3 in Hanasusanto and Kuhn (2017))**. *If there exists a real vector $\boldsymbol{z}$ such that $A_1^{1^\top} z > 0$, then $A_1^{1^\top} A_1^1 \succ_{co} 0$.*

Given this result, suppose for an A_1^1, there does not exist such a vector $\boldsymbol{z}$. It implies A_1^1 does not have full row rank. In this case, we can add rows in A_1^1 to increase the row rank. As long as the added rows constitute valid cuts to the problem (1), we will have a valid completely positive program formulation. In the next section, we demonstrate how this approach can be used to derive strong duality result for the risk management problem we studied in this paper.

4. Risk Mitigation Models

In this section, we solve a two-stage distributionally robust problem. In the first stage, given reported TTRs by each vendor, we obtain the optimal inventory levels at each node, so that in the second stage (after disruption events happen), there are no or minimal lost sales during the disruption period.

Network Structure:
We consider a supply chain network consisting p vendor nodes (in set $\mathcal{P}$) and n customer nodes (in set $\mathcal{N}$). For those plants supplying multiple products, we split them into multiple nodes such that each node represents one type of product. For instance, supposing vendor node $k \in \mathcal{P}$ produces x different products, we replace vendor k in the network by x "plant nodes" (grouped into set $\mathcal{A}_k$) and obtain a new supply chain graph. Figure 4 gives a simple example. Vendor 1 in this example



(a) Original supply chain (b) Rebuilt network

Figure 4 Example for a rebuilt network

produces both steel and copper. We split it into two nodes with one node for steel production and the other for copper production. We refer to the nodes before splitting as "vendor nodes" and those nodes after splitting as "plant nodes" or simply "nodes". In this example, Vendor node 1 is split in to Plant node 1 and Plant node 2. Vendor node 2 only produces rubber and is not split. It is then also labeled as Plant node 3. After rebuilding the supply chain in this way, we now have a new network, $\mathcal{G}$, containing an enlarged number of plant nodes, say m plant nodes (in set $\mathcal{M}$). We partition those nodes in $\mathcal{M}$ into set $\mathcal{M}_1$ if the production requires no input from other nodes, and $\mathcal{M}_2$ otherwise. Let $\mathcal{T}^p$ denote the set of all product types; $\mathcal{T}_j$ denote the set of raw material types needed by plant node $j, \forall j \in \mathcal{M}_2$ (We do not consider the raw material flows to the highest tier suppliers). Let $t_p = |\mathcal{T}^p|$. For ease of reading, the notations, parameters and assumptions used in our model are summarized in Table 1.

Table 1 Notations, parameter, and assumptions

		Notation	Definition	Assumptions
The first-stage inventory problem	Decision variables	$r_i \in \mathbb{R}_+, i \in \mathcal{M}$	The finished good inventory in plant i	We assume plant nodes hold finished good inventory instead of unprocessed parts.
	Parameters	$h_i \in \mathbb{R}_+, i \in \mathcal{M}$	Unit inventory holding cost for plant node i.	
		$br \in \mathbb{R}_+$	The total inventory budget for the whole supply chain.	
The second-stage production problem	Decision variables	$x_{ij} \in \mathbb{R}_+, (i,j) \in \mathcal{G}$	the material flow from plant node i to plant node j.	
		$u_i \in \mathbb{R}_+, i \in \mathcal{M}$	The number of goods produced at plant node i.	
		$l_i \in \mathbb{R}_+, i \in \mathcal{N}$	The lost sales at customer node i.	
	Parameters	$d_i \in \mathbb{R}_+, i \in \mathcal{N}$	The demand rate at customer node i.	We assume demand rates are deterministic.
		$f_i \in \mathbb{R}_+, i \in \mathcal{N}$	The penalty cost for one unit of lost sales at customer node i.	
		$c_i \in \mathbb{R}_+, i \in \mathcal{P}$	Production capacity per unit time of vendor node i.	We assume producing one unit of finished goods (regardless of the types of finished goods) consumes one unit of capacity.
		$B \in \mathbb{R}_+^{tp \times m}$	BOM matrix, with B_{tj} denoting number of type t product needed to produce 1 unit of item in plant node j.	
		$I^{PT} \in \{0,1\}^{m \times tp}$	The indicator matrix with the entry on row i and column t equal to 1 if the product produced by plant node i is of type t.	
Disruption risk and duration	Random variables	$v_i \in \{0,1\}, i \in \mathcal{P}$	the survival indicator for vendor node i,,i.e. if vendor i is not disrupted, $v_i = 1$, otherwise, $v_i = 0$.	We assume once vendor i is disrupted, and all the plant nodes split from vendor node i cannot produce anything.
		$T^R(\boldsymbol{v}) \in \mathbb{R}_+$	The supply chain's TTR. The value of supply chain TTR depends on disruption scenarios. So we represent it as a function of survival indicator.	
	Parameters	$T_i^r \in \mathbb{R}_+, i \in \mathcal{P}$	The time-to-recover (TTR) reported by vendor node i;	We assume each vendor's TTR is deterministic.

4.1. Supply Chain Resilience under Disruption

In general, there are many ways to measure the resiliency of a supply chain. We focus first on the performance of lost sales in this paper, with deterministic TTRs of vendors, and discuss later how this methodology can be used for other classes of disruption problems under more general settings. More specifically, for a given disruption scenario $\boldsymbol{v}$ and strategic inventory deployment $\boldsymbol{r}$, the optimal recovery operation to minimize total cost of lost sales during the disruptions can be modelled as an LP:

$$
\begin{array}{llll}
Z(\boldsymbol{v},\boldsymbol{r}) = & \min\limits_{(x_{ij},\boldsymbol{u},\boldsymbol{l})} & \sum\limits_{j\in\mathcal{N}} f_j l_j & \\
& s.t. & \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij} + l_j \geq d_j T^R(\boldsymbol{v}), & \forall j\in\mathcal{N} \\
& & \sum\limits_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij} - u_i \leq r_i, & \forall i\in\mathcal{M} \\
& & \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij}I_{it}^{PT}}{B_{tj}} - u_j \geq 0, & \forall j\in\mathcal{M}_2, t\in\mathcal{T}_j \\
& & \sum\limits_{i\in\mathcal{A}_k} u_i \leq (T^R(\boldsymbol{v}) - T_k^r(1-v_k))c_k, & \forall k\in\mathcal{P} \\
& & x_{ij}\geq 0, \boldsymbol{u},\boldsymbol{l}\geq \boldsymbol{0} &
\end{array}
\tag{6}
$$

Note that $T^R(\boldsymbol{v})$ denotes the supply chain's Time-To-Recovery, and $(T^R(\boldsymbol{v})-T_k^r(1-v_k))c_k$ denotes the total production capacity available during the disruption duration, after accounting for the lost production if vendor node k is disrupted. The first constraint defines the lost sales during supply chain TTR for each final goods. The second constraint specifies the total outflow of a plant must be bounded by units produced and inventory held. The third constraint means the total production in a plant is constrained by the raw materials supplied from upstream. Finally, the last constraint indicates the total units produced in all plant nodes split from a single vendor node cannot exceed available capacity during supply chain's TTR of this vendor node.

There are numerous ways to model the supply chain TTR (i.e., $T^R(\boldsymbol{v})$), as a function of the disruption indicated by the random variables $\boldsymbol{v}$. For instance, we could assume that all disruption to the supply chain happens at the same time, i.e., $T^R(\boldsymbol{v}) = \max_k(T_k^r(1-v_k))$, or model the cascading effect of allowing one facility to fail due to disruption at a nearby facility, after some random duration. In this way, we need to model $T^R(\boldsymbol{v})$ using a more refined stochastic model. In our distributionally robust approach, we assume that only the means and covariances of $\boldsymbol{v}$ and $T^R(\boldsymbol{v})$ are known, and enforce only the weaker constraint

$$T^R(\boldsymbol{v}) \geq T_k^r(1-v_k),\ \forall\ k.$$

To rule out pathological cases, we assume further that the system is carefully configured such that all production and replenishment activities in the supply chain are needed during normal operation. To enforce this, we assume that when the production capacity is properly scaled by a factor of $1+\epsilon$ for some $\epsilon>0$, all flows and production nodes can work at a positive rate during normal operation. In other words, there exists strictly positive x_{ij}^0 and $\boldsymbol{u}^0$ such that

$$
\left\{
\begin{array}{ll}
\sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij}^0 > d_j, & \forall j\in\mathcal{N} \\
-\sum\limits_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij}^0 + u_i^0 > 0, & \forall i\in\mathcal{M} \\
\sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij}^0 I_{it}^{PT}}{B_{tj}} - u_j^0 > 0, & \forall j\in\mathcal{M}_2, t\in\mathcal{T}_j \\
\sum\limits_{i\in\mathcal{A}_k} u_i^0 < (1+\epsilon)c_k, & \forall k\in\mathcal{P} \\
x_{ij}^0 > 0, \boldsymbol{u}^0 > \boldsymbol{0} &
\end{array}
\right\}
\tag{7}
$$

It turns out that this assumption is crucial for strong duality to hold in our conic programming reformulation of this problem.

We show next how some of the more restrictive assumptions used in the model development can be removed, to address more general disruption mitigation problem.

4.1.1. Random Time-To-Recovery. Suppose TTR for each vendor is not deterministic but follows a discrete distribution, says $T_i^r = T_i(k)$ with probability $p_i(k)$, $k = 0, \ldots, K$, with $T_i(0) = 0$ indicating no disruption.

We duplicate the vendor node i into $K+1$ copies, each with capacity c_i, but with TTR equals $T_i(k)$ for the kth copy. The disruption event for each node satisfies: $v_i(k) \in \{0,1\}, \sum_k (1 - v_i(k)) = 1$, and each $v_i(k) = 0$ with probability $p_i(k)$.

We link the kth copy to the $(k+1)$th copy, replacing the vendor node i in the network with a serial graph with $K+1$ nodes. All arcs entering vendor node i now enter via the first copy ($k = 0$), and all arcs leaving vendor node i now leave via the Kth copy. In this way, we ensure that the total production capacity coming from this serial graph is not more than $\max(T^R(v) - \max_k(T_i(k)(1 - v_i(k)))c_i$.

We put all strategic inventory at the Kth copy, to ensure that the inventory will not be destroyed due to disruption at vendor node i. This transformation ensures that we can convert the random TTR problems into one with deterministic TTR, and solved using the model proposed earlier.

4.1.2. Resiliency using general measurement. Note that the performance of any risk mitigation strategy depend on the following key parameters:

- $T^R(\boldsymbol{v})$: The duration of the supply chain disruption;
- $T_k^r(1 - v_k)c_k$: The amount of production affected due to disruption; and
- r_i: Inventory deployed at plant node i in anticipation of the disruption.

While we restrict our discussion to the model used in the Risk-Exposure-Index literature, our approach is applicable as long as the resiliency measurement $Z(\boldsymbol{v}, \boldsymbol{r})$ is a piecewise linear convex in the variables $T^R(\boldsymbol{v}), \boldsymbol{r}$, and $T^R(\boldsymbol{v}) - T_k^r(1 - v_k)$ for each vendor node k in the network.

For example, we can interpret the penalty cost f_j as emergency shipment cost for unmet demand, and change the objective function in (6) to $\sum_j f_j l_j + \sum_{i,j} c_{i,j} x_{i,j}$, where $c_{i,j}$ is unit shipment cost from i to j, to measure resiliency of the supply chain using the notion of 'total shipment cost under disruption". Note that the objective function is a piecewise linear convex function in the variables $T^R(\boldsymbol{v}), \boldsymbol{r}$, and $T^R(\boldsymbol{v}) - T_k^r(1 - v_k)$. The co-positive cone framework can be easily extended to this case.

4.2. Inventory Mitigation Model - Distributionally Robust Approach

We assume that the disruption distribution information is partially revealed to the decision maker. Specifically, survival indicator $v_i, i = 1, ..., p$ are random binary variables. We assume the set of distribution of $\tilde{\boldsymbol{v}} \sim (\mathcal{D}(\mathcal{B}), \boldsymbol{\mu}^v, \Sigma^v)$, is defined by the binary support $\{0,1\}^p_+$, with finite mean $\boldsymbol{\mu^v}$ and second-moment matrix Σ^v. We use $\tilde{v}_i$ to indicate v_i here is a random variable. To get the first two moments of $T^R(\boldsymbol{v})$, we can simulate the disruptions according to moments of $\tilde{\boldsymbol{v}}$. By the definition of $T^R(\boldsymbol{v})$, we can summarize the mean and standard deviation from the simulation results. We denote the mean and the second moment as μ_T, σ_T respectively.

The objective is to explore the optimal inventory allocations so that the WCVaR of total lost sales during disruption period is within a threshold, say c_0. Specifically, we consider the following problem:
Given confidence level $1 - \eta$,

$$\begin{aligned} &\min_{\boldsymbol{r} \geq 0} \ \boldsymbol{h}^\top \boldsymbol{r} \\ &s.t. \ \ WCVaR_{1-\eta} = \min_{\theta} \left\{ \theta + \tfrac{1}{\eta} \max\nolimits_{\tilde{v} \sim (\mathcal{D}(\mathcal{B}), \boldsymbol{\mu}^v, \Sigma^v), T^R(\boldsymbol{v}) \sim (\mathbb{R}_+, \mu_T, \sigma_T)} \mathbf{E}[(Z(\tilde{\boldsymbol{v}}, T^R(\tilde{\boldsymbol{v}}), r) - \theta)^+] \right\} \leq c_0 \end{aligned} \tag{8}$$

or equivalently,

$$\begin{aligned} &\min_{\boldsymbol{r} \geq \mathbf{0}, \ \theta} \ \boldsymbol{h}^\top \boldsymbol{r} \\ &s.t. \quad \theta + \tfrac{1}{\eta} \max\nolimits_{\tilde{v} \sim (\mathcal{D}(\mathcal{B}), \boldsymbol{\mu}^v, \Sigma^v), T^R(\boldsymbol{v}) \sim (\mathbb{R}_+, \mu_T, \sigma_T)} \mathbf{E}[(Z(\tilde{\boldsymbol{v}}, T^R(\tilde{\boldsymbol{v}}), r) - \theta)^+] \leq c_0 \end{aligned} \tag{9}$$

where $Z(\tilde{\boldsymbol{v}}, T^R(\tilde{\boldsymbol{v}}), r)$ denotes the minimum lost sales given disruption scenario $\tilde{\boldsymbol{v}}$ and inventory deployment $\boldsymbol{r}$.

Alternately, we can minimize the CVaR of lost sales, with given inventory budget br,using the following formulation:

$$\begin{aligned} &\min_{\boldsymbol{r} \geq \mathbf{0}} \ \theta + \tfrac{1}{\eta} \max\nolimits_{\tilde{v} \sim (\mathcal{D}(\mathcal{B}), \boldsymbol{\mu}^v, \Sigma^v), T^R(\boldsymbol{v}) \sim (\mathbb{R}_+, \mu_T, \sigma_T)} \mathbf{E}[(Z(\tilde{\boldsymbol{v}}, T^R(\tilde{\boldsymbol{v}}), r) - \theta)^+] \\ &s.t. \ \ \boldsymbol{h}^\top \boldsymbol{r} \leq br \end{aligned} \tag{10}$$

4.3. Benchmark Inventory Mitigation Model 1 - The Traditional REI Approach

We use the REI approach as the first benchmark for our analysis. This builds on the standard stochastic programming methodology, using scenario decomposition to synthesize the impact of the mitigation strategy in each scenario. Furthermore, we assume that in each scenario only one vendor node can be disrupted (i.e., all the corresponding plant nodes are disrupted). Each scenario w is indexed by elements in $\{1, 2, \ldots, p\}$, and $T^{R^{(w)}}$ denotes the supply chain TTR in scenario w.

To ensure that we have sufficient inventory in the system to prevent lost sales in all scenarios, we solve a related LP in (11).

$$
\begin{array}{llll}
\min \boldsymbol{h}^{\intercal}\boldsymbol{r} & & & \\
s.t. & \sum\limits_{i\in\mathcal{M}:(i,j)\in\mathcal{G}} x_{ij}^{(w)} & \geq d_j T^{R^{(w)}}, & \forall j\in\mathcal{N}, w\in\{1,2,..,p\} \\
& \sum\limits_{j\in\mathcal{M}\cup\mathcal{N}:(i,j)\in\mathcal{G}} x_{ij}^{(w)} - u_i^{(w)} & \leq r_i, & \forall i\in\mathcal{M}, w\in\{1,2,..,p\} \\
& \sum\limits_{i\in\mathcal{M}:(i,j)\in\mathcal{G}} \frac{x_{ij}^{(w)} I_{it}^{PT}}{B_{tj}} - u_j^{(w)} & \geq 0, & \forall j\in\mathcal{M}_2, t\in\mathcal{T}_j, w\in\{1,2,..,p\} \\
& \sum\limits_{i\in A_k} u_i^{(w)} & \leq (T^{R^{(w)}} - T_k^{r}(1-v_k^{(w)}))c_k, & \forall k\in\mathcal{P}, w\in\{1,2,..,p\} \\
& v_w^{(w)} & = 1 & \forall w\in\{1,2,..,p\} \\
& v_k^{(w)} & = 0 & \forall k\neq w, \forall w\in\{1,2,..,p\} \\
& x_{ij}^{(w)}\geq 0, u_j^{(w)}\geq 0, & & \forall w\in\{1,2,..,p\} \\
& \boldsymbol{r}\geq\boldsymbol{0} & &
\end{array}
\tag{11}
$$

Note that $v_k^{(w)}$ denotes the event that vendor node k is operational in scenario w, and $T^{R^{(w)}} = \max_k T_k^{r}(1-v_k^{(w)})$ as there is only one disruption in each scenario. This model can be used to find an inventory allocation strategy $\boldsymbol{r}$ with the smallest total investment, so that the supply chain always has zero lost sales when at most one vendor is disrupted. The first constraint guarantees no lost sales are incurred in each demand node for all p scenarios. The second and third constraints are flow conservation constraints based on the inventory available. The fourth constraint indicates for each vendor, its total production units are bounded by its capacity during recovered time period. The next two constraints give corresponding disruption indicators' values for scenario w.

4.4. Benchmark Inventory Mitigation Model 2 - The Stochastic Programming Approach

In another extreme, we assume the disruption probability distribution is fully known to the decision maker, and the risk measure adopted is $CVaR$. Specifically, let $p^{(w)}, w\in S$, denote the probability of disruption scenario w occurring, where S is the set of all disruption scenarios. We would like to solve the following stochastic program.

$$
\begin{array}{ll}
\min\limits_{\boldsymbol{r}\geq\boldsymbol{0},\theta} & \theta + \frac{1}{\eta}\sum\limits_{w\in S} p^{(w)}[(Z_{(w)}(\boldsymbol{r})-\theta)^{+}] \\
s.t. & \boldsymbol{h}^{\intercal}\boldsymbol{r}\leq br
\end{array}
\tag{12}
$$

Where $Z_{(w)}$ is the total lost sales under disruption scenario w. Note that we fix the total inventory budget in this formulation, rather than minimizing the total inventory cost as in the first REI benchmark case. This is to facilitate the comparison among the REI model, the stochastic model and the distributionally robust model. Specifically, we will first solve REI problem (11) to obtain the minimum total inventory cost. Then we use the minimum total inventory cost value as a fixed total budget for both stochastic model and our distributionally robust model to obtain the respective inventory strategies so that all three strategies are with the same total inventory cost.

Note that Problem (12) can be equivalently reformulated as the following linear program.

$$
\begin{aligned}
&\min\ \theta+\tfrac{1}{\eta}\sum_{w\in S}p^{(w)}Q^{(w)}(\boldsymbol{r},\theta)\\
&s.t.\ \ \boldsymbol{h}^{\mathsf{T}}\boldsymbol{r}\le br\\
&\qquad\ \boldsymbol{r}\ge\boldsymbol{0}
\end{aligned}
\tag{13}
$$

where for any $w=1,...,|S|$,

$$
\begin{aligned}
Q^{(w)}(\boldsymbol{r},\theta)=\ &\min\ y^{(w)}\\
&s.t.\ \ y^{(w)}\ge\sum_{j\in\mathcal{N}}f_j^{(w)}l_j^{(w)}-\theta\\
&\qquad\sum_{i\in\mathcal{M}:(i,j)\in\mathcal{G}}x_{ij}^{(w)}+l_j^{(w)}\ge d_jT^{R(w)}, &&\forall j\in\mathcal{N}\\
&\qquad\sum_{j\in\mathcal{M}\cup\mathcal{N}:(i,j)\in\mathcal{G}}x_{ij}^{(w)}-u_i^{(w)}\le r_i, &&\forall i\in\mathcal{M}\\
&\qquad\sum_{i\in\mathcal{M}:(i,j)\in\mathcal{G}}\frac{x_{ij}^{(w)}I_{it}^{PT}}{B_{tj}}-u_j^{(w)}\ge 0, &&\forall j\in\mathcal{M}_2,t\in\mathcal{T}_j\\
&\qquad\sum_{i\in A_k}u_i^{(w)}\le(T^{R(w)}-T_k^r(1-v_k^{(w)}))c_k, &&\forall k\in\mathcal{P}\\
&\qquad y^{(w)}\ge 0\\
&\qquad x_{ij}^{(w)},u_j^{(w)},l_j^{(w)}\ge 0,\forall i\in\mathcal{M},j\in\mathcal{N}
\end{aligned}
\tag{14}
$$

Classical Benders decomposition technique can be used to find an approximate solution to this large scale LP problem.

4.5. Distributionally Robust Inventory Mitigation Model

Consider the lost sales problem (6). Let $\boldsymbol{\alpha}\in\mathbb{R}_+^{|\mathcal{N}|},\boldsymbol{\beta}\in\mathbb{R}_+^{|\mathcal{M}|},\boldsymbol{\gamma}\in\mathbb{R}_+^{tp},\boldsymbol{\delta}\in\mathbb{R}_+^{|\mathcal{P}|}$ denote the dual variables corresponding to each set of constraints in Problem (6). $\boldsymbol{s^1},\boldsymbol{s^2},\boldsymbol{s^3},\boldsymbol{s^4},\boldsymbol{s^5}$ are slack variables. Due to strong duality property of linear program, we have an equivalent dual formulation as follows.

$$
\begin{aligned}
&\max\ \sum_{j\in\mathcal{N}}d_j\alpha_jT^R(\boldsymbol{v})-\sum_{i\in\mathcal{M}}r_i\beta_i-\sum_{k\in\mathcal{P}}c_k\delta_k(T^R-T_k^r(1-v_k))\\
&s.t.\ \ (\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{s^1},\boldsymbol{s^2},\boldsymbol{s^3},\boldsymbol{s^4},\boldsymbol{s^5})\in\mathcal{F}
\end{aligned}
\tag{15}
$$

where $\mathcal{F}$ is the feasible polyhedron of dual problem.

$$
\mathcal{F}=\left\{
\begin{array}{ll}
\alpha_j-\beta_i+s_l^1=0, & \forall j\in\mathcal{N},(i,j)\in\mathcal{G}_1\\
-\beta_i+\sum_{t\in\mathcal{T}_j}\frac{\gamma_t^jI_{it}^{PT}}{B_{tj}}+s_l^2=0, & \forall i\in\mathcal{M},\forall j\in\mathcal{M}_2,(i,j)\in\mathcal{G}\\
-\delta_k+\beta_{i(k)}-\sum_{t\in\mathcal{T}_{i(k)}}\gamma_t^{i(k)}+s_l^3=0, & \forall k\in\mathcal{P},i(k)\in\mathcal{A}_k,i(k)\in\mathcal{M}_2\\
-\delta_k+\beta_{i(k)}+s_l^4=0, & \forall k\in\mathcal{P},i(k)\in\mathcal{A}_k,i(k)\in\mathcal{M}_1\\
\alpha_j+s_j^5=f_j & \forall j\in\mathcal{N}\\
\boldsymbol{\alpha}\in\mathbb{R}_+^n,\boldsymbol{\beta}\in\mathbb{R}_+^m,\boldsymbol{\gamma}\in\mathbb{R}_+^{tp},\boldsymbol{\delta}\in\mathbb{R}_+^p & \\
\boldsymbol{s^1}\in\mathbb{R}_+^{|\mathcal{G}_2|},\boldsymbol{s^2}\in\mathbb{R}_+^{|\mathcal{G}_1|},\boldsymbol{s^3}\in\mathbb{R}_+^{m_2} & \\
\boldsymbol{s^4}\in\mathbb{R}_+^{m_1},\boldsymbol{s^5}\in\mathbb{R}_+^n &
\end{array}
\right\}
$$

The nonsmooth nature of CVaR and WCVaR poses a challenge in modeling and numerical computation. We overcome it by reformulating $(Z(\tilde{\boldsymbol{v}}, T^R(\tilde{\boldsymbol{v}}), r) - \theta)^+$ into the following:

$$
\begin{aligned}
(Z(\boldsymbol{v}, T^R(\tilde{\boldsymbol{v}}), \boldsymbol{r}) - \theta)^+ = \max\ & (\sum_{j\in\mathcal{N}} d_j\alpha_j - \sum_{k\in\mathcal{P}} c_k\delta_k)T^R - \sum_{i\in\mathcal{M}} r_i\beta_i + \sum_{k\in\mathcal{P}} c_k\delta_k T_k^r(1-\tilde{v}_k) - \theta y \\
s.t.\ & (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{s^1}, \boldsymbol{s^2}, \boldsymbol{s^3}, \boldsymbol{s^4}, \boldsymbol{s^5}) \in \mathcal{F} \\
& (1-y)(\sum_{j\in\mathcal{N}} \alpha_j + \sum_{i\in\mathcal{M}} \beta_i + \sum_{k\in\mathcal{P}} \delta_k) = 0 \\
& y \in \{0, 1\}
\end{aligned} \tag{16}
$$

The introduction of binary variable y guarantees $(Z(\boldsymbol{v}, T^R(\tilde{\boldsymbol{v}}), r) - \theta)^+$ takes value of $Z(\boldsymbol{v}, T^R(\tilde{\boldsymbol{v}}), r) - \theta$ when it is greater than 0, and 0 otherwise. The last quadratic constraint ensures that the dual variables take value 0 when $y \neq 1$, since they are assumed to be nonnegative.

To make the formula more compact, we can add some valid cuts in the model. First we observe

Lemma 3. *The optimal dual variables* $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{\gamma}^*, \boldsymbol{\delta}^*)$ *in Problem (15) satisfies:* $\sum_{j\in\mathcal{N}} d_j\alpha_j^* \leq \sum_{k\in\mathcal{P}} c_k\delta_k^*$

Lemma 3 indicates that we can add the cut $\sum_{j\in\mathcal{N}} d_j\alpha_j \leq \sum_{k\in\mathcal{P}} c_k\delta_k$ in Problem (15) without cutting off the optimal solutions. We introduce a slack variable s^7 to make this cut a equality form.

Notice that for assumption (A1) in Section 3 to hold, the feasible region has to be bounded. By analysing the structure of feasible region $\mathcal{F}$, we add several valid cuts ((16a) to (16f) given in Lemma 4) to Problem (16) to bound the feasible region.

Lemma 4. *We have following constraints as valid cuts to the problem.*

$$\beta_j + s_j^8 = \left(\max_{i=1}^{n} f_i\right) y, \forall j = 1, \ldots, m \tag{17a}$$

$$\delta_k + s_k^9 = \left(\max_{i=1}^{n} f_i\right) y, \forall k = 1, \ldots, p \tag{17b}$$

$$y + s^{10} = 1 \tag{17c}$$

$$\alpha_i s_i^5 = 0, \forall i = 1, \ldots, n \tag{17d}$$

$$\alpha_j + s_j^{11} = f_j y \tag{17e}$$

$$\boldsymbol{s^8}, \boldsymbol{s^9}, s^{10}, \boldsymbol{s^{11}} \geq \mathbf{0} \tag{17f}$$

According to the definition of $T^R(\boldsymbol{v})$, $T^R(\boldsymbol{v}) \geq T_k^r(1 - v_k), \forall k \in \mathcal{P}$. Hence we can also add this constraint to the problem by introducing slack variable $\boldsymbol{s}^6$ as follows:

$$T^R(\boldsymbol{v}) - s_k^6 = T_k^r(1 - v_k), \forall k \in \mathcal{P} \tag{18a}$$

$$\boldsymbol{s^6} \geq 0 \tag{18b}$$

Let $Z^m(\boldsymbol{r},\theta)$ denote $\max_{\tilde{v}\sim(\mathcal{D}(\mathcal{B}),\boldsymbol{\mu}^v,\Sigma^v),T^R(\boldsymbol{v})\sim(\mathbb{R}_+,\mu_T,\sigma_T)} \mathbf{E}[(Z(\tilde{\boldsymbol{v}},T^R(\tilde{\boldsymbol{v}}),r)-\theta)^+]$. We can reformulate the problem as:

$$
Z^m(\boldsymbol{r},\theta) = \max_{\tilde{v}\sim(\mathcal{D}(\mathcal{B}),\boldsymbol{\mu}^v,\Sigma^v),T^R(\boldsymbol{v})\sim(\mathbb{R}_+,\mu_T,\sigma_T)} \mathbf{E}\left[\begin{array}{ll} \max & \sum_{j\in\mathcal{N}} d_j\alpha_j T^R(\boldsymbol{v}) - \sum_{i\in\mathcal{M}} r_i\beta_i - \sum_{k\in\mathcal{P}} c_k\delta_k s_k^6 - \theta y \\ s.t. & (\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{s^1},\boldsymbol{s^2},\boldsymbol{s^3},\boldsymbol{s^4},\boldsymbol{s^5}) \in \mathcal{F} \\ & (1-y)(\sum_{j\in\mathcal{N}}\alpha_j + \sum_{i\in\mathcal{M}}\beta_i + \sum_{k\in\mathcal{P}}\delta_k) = 0 \\ & y\in\{0,1\} \\ & \sum_{j\in\mathcal{N}} d_j\alpha_j + s^7 = \sum_{k\in\mathcal{P}} c_k\delta_k \\ & (16a)-(16f),(17a)-(17b) \end{array}\right] \tag{19}
$$

It turns out that these valid cuts are extremely important to the approach used, since we can only obtain good numerical performance after adding these cuts. Our theoretical analysis also shows that the conic program has zero duality gap if these cuts are added into the model. To see it, noticed that Problem (19) falls in the realm of framework considered in Proposition 3. Specifically, we decompose the decision variables into two parts: $\boldsymbol{x}_2$ represents $\boldsymbol{s}^6$ in (17a) and the slack variables $\boldsymbol{s}$ in the valid moment matrix, $\mathcal{M}(\mathcal{B})$; $\boldsymbol{x}_1$ includes $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta}$ and $\boldsymbol{s}^i$, for $i\neq 6$. With this decomposition, Problem (19) satisfies problem structure requirements (i) to (iii) in Proposition 3. Moreover, $A_2^2 = \begin{pmatrix} -I & O \\ O & I \end{pmatrix}$, which meets the condition $A_1^{2\mathsf{T}}A_1^2 + A_2^{2\mathsf{T}}A_2^2 \succ 0$. Therefore, to show the strong duality, it is sufficient to show $A_1^{1\mathsf{T}}A_1^1 \succ_{co} 0$.

LEMMA 5. *If there exists strictly positive solution to feasible region defined by (7), A_1^1 defined in (19) satisfies $A_1^{1\mathsf{T}}A_1^1 \succ_{co} 0$.*

According to Proposition 3, there is therefore no duality gap between the completely positive formulation and co-positive formulation of $Z^m(\boldsymbol{r},\theta)$.

4.6. First-Stage Co-positive Formulation

With the dual formulation to the second stage problem which is denoted as Z^m_{co}, we can convert the following two stage problem

$$
\begin{array}{ll} \min_{\boldsymbol{r}\geq\boldsymbol{0},\theta} & \boldsymbol{h}^\mathsf{T}\boldsymbol{r} \\ s.t. & \left\{\theta + \frac{1}{\eta}Z^m_{co}(\boldsymbol{r},\theta)\right\} \leq c_0 \end{array} \tag{20}
$$

into an equivalent co-positive cone problem

$$
\begin{array}{ll} \min_{\boldsymbol{r}\geq\boldsymbol{0}} & \boldsymbol{h}^\mathsf{T}\boldsymbol{r} \\ s.t. & \theta + \frac{1}{\eta}(\rho + \boldsymbol{\mu}^{v\mathsf{T}}\boldsymbol{\nu} + \Sigma^v \bullet \Theta + \boldsymbol{b}_1^\mathsf{T}\boldsymbol{\phi}_x + \boldsymbol{b}_1^\mathsf{T}\Lambda(\boldsymbol{\epsilon}_x)\boldsymbol{b}_1 + \mathbf{1}_p^\mathsf{T}\boldsymbol{\phi}_{\boldsymbol{w}} + \mathbf{1}_p^\mathsf{T}\epsilon_w\mathbf{1}_p) \leq c_0 \\ & CO \succcurlyeq_{co} 0 \end{array} \tag{21}
$$

where CO is the dual co-positive matrix. We omit the explicit expression for simplicity.

4.7. Fixed Inventory Budget

Recall that when there is a budget on the total holding cost, the optimal inventory allocation for a fixed inventory budget can be formulated as Problem (10). By exactly the same derivation, we can obtain its corresponding co-positive program presented below.

$$\begin{array}{lc} \min_{\boldsymbol{r}\geq \mathbf{0}} \ \theta + \frac{1}{\eta}(\rho + \boldsymbol{\mu}^{v\intercal}\boldsymbol{\nu} + \Sigma^v \bullet \Theta + \boldsymbol{b}_1^\intercal \boldsymbol{\phi}_x + \boldsymbol{b}_1^\intercal \Lambda(\boldsymbol{\epsilon}_x)\boldsymbol{b}_1 + \mathbf{1}_p^\intercal \boldsymbol{\phi}_{\boldsymbol{w}} + \mathbf{1}_p^\intercal \epsilon_w \mathbf{1}_p) & \\ s.t. \qquad \boldsymbol{h}^\intercal \boldsymbol{r} \leq br & \\ \qquad\qquad CO \succcurlyeq_{co} 0 & \end{array} \tag{22}$$

4.8. The Worst-Case Expected Lost Sales

For the sake of completeness, we next show that the worst-case "expected lost sale" under optimal inventory allocation, $Z^{\hat{e}}(\boldsymbol{r}^*)$, can be solved as a special case of the worst case CVaR.

Lemma 6. *Given the optimal inventory allocation $\boldsymbol{r}^*$, let $Z^{\hat{e}}(\boldsymbol{r}^*)$ be the worst-case expected lost sales, i.e.*

$$Z^{\hat{e}}(\boldsymbol{r}^*) := \max_{\tilde{v}\sim(\mathcal{D}(\mathcal{B}),\boldsymbol{\mu}^v,\Sigma^v),T^R(\boldsymbol{v})\sim(\mathbb{R}_+,\mu_T,\sigma_T)} \boldsymbol{E}[Z(\tilde{\boldsymbol{v}},\boldsymbol{r}^*)] \tag{23}$$

Then $Z^{\hat{e}}(\boldsymbol{r}^*)$ can be solved by the completely positive reformulation of the worst-case CVaR by setting $\boldsymbol{r} = \boldsymbol{r}^*$ and $\theta = 0$, i.e. $Z^{\hat{e}}(\boldsymbol{r}^*) = Z^m_{cp}(\boldsymbol{r}^*, 0)$.

5. Sensitivity Analysis

In the previous section, we have obtained a characterization of the worst-case solution to a distributionally robust supply chain disruption problem, using recent results developed in the theory of conic programming. While there are various ways to formulate such problems in a robust manner, our approach has the advantage that it admits a probabilistic interpretation in terms of the worst-case distribution to the robust problem, and uses directly the parsimonious set of risk estimates often encountered in practice. This approach can be used to derive insightful information on the sensitivity analysis of the key parameters used in the model. We provide the intuition behind the sensitivity results using the worst-case distribution interpretation, and relegate the rigorous proof utilizing the conic program to the online companion.

5.1. Impact of Supplier's TTR

Note that in our risk mitigation model, the TTR values for the vendors are self reported. The risk mitigation strategy is predicated on the assumption that these values reflect accurately vendors' ability to recover from disruption. This assumption requires the vendor to have a reliable procedure to estimate the time duration within which it can recover from a certain type of disruption. What happen if suppliers can be nudged to reduce their TTRs? Which vendor should we focus on to reduce the TTR? In this section, we apply sensitivity analysis on the reported TTR to see how the changes in supplier's TTR affect the worst-case expected lost sales. Note that our analysis focused on sensitivity analysis in the worst-case scenario, instead of the more traditional expected lost sales setting.

Recall that for fixed $\boldsymbol{v}$ and $\boldsymbol{r}$, the lost sales are given by

$$\begin{array}{llll}
Z(\boldsymbol{v},\boldsymbol{r}) = \min\limits_{(x_{ij},\boldsymbol{u},\boldsymbol{l})} & \sum\limits_{j\in\mathcal{N}} f_j l_j & & \text{Dual Variables} \\
s.t. & \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij} + l_j \geq d_j T^R(\boldsymbol{v}), & \forall j\in\mathcal{N} & \boldsymbol{\alpha} \\
& \sum\limits_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij} - u_i \leq r_i, & \forall i\in\mathcal{M} & \boldsymbol{\beta} \\
& \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij} I_{it}^{PT}}{B_{tj}} - u_j \geq 0, & \forall j\in\mathcal{M}, t\in\mathcal{T}_j & \boldsymbol{\gamma} \\
& \sum\limits_{i\in\mathcal{A}_k} u_i \leq (T^R(\boldsymbol{v}) - T_k^r(1-v_k))c_k, & \forall k\in\mathcal{P} & \boldsymbol{\delta} \\
& x_{ij}\geq 0, \boldsymbol{u},\boldsymbol{l}\geq \boldsymbol{0} & &
\end{array}$$

This is an LP, and the optimal dual solution $\boldsymbol{\delta}$ provides valuable information on the sensitivity of the input T_k^r for our problem. In particular, we know that in the optimal solution,

- If $v_k = 1$, then change in T_k^r does not affect the optimal solution $Z(\boldsymbol{v},\boldsymbol{r})$.
- If $v_k = 0$ and $T_k^r < T^R(\boldsymbol{v})$, then a unit decrease in T_k^r resulted in a decrease of $c_k\delta_k(\boldsymbol{v})$, where $\delta_k(\boldsymbol{v})$ is the corresponding dual solution.
- The situation when $v_k = 0$ and $T_k^r = T^R(\boldsymbol{v})$ is more complicated, and depends on whether there are multiple facilities attaining the same $T^R(\boldsymbol{v})$ in the system. In the case that vendor k attains the maximum TTR alone, (and assuming after a unit decrease, k's TTR is still the largest among all,) then a unit decrease in T_k^r will result in a change of $-\sum_{j\in\mathcal{N}}\alpha_j(\boldsymbol{v})d_j + \sum_{i\neq k} c_i\delta_i(\boldsymbol{v})$ in lost sales. This change is therefore at least $-c_k\delta_k(\boldsymbol{v})$ according to Lemma 3. On the other hand, when k attains the maximum TTR along with other disrupted vendor nodes, a decrease of a unit in T_k^r will result in a decrease of $c_k\delta_k(\boldsymbol{v})$ units in lost sales, as in the previous case.

In summary, we expect the function $c_k\delta_k(\boldsymbol{v})(1-v_k)$ to be a upper bound on the lost sales when T_k^r decreases by a unit. Since $\tilde{\boldsymbol{v}}$ is random, we expect the function $c_k\mathbf{E}[\delta_k(\boldsymbol{v})(1-v_k)]$ to be an upper bound for the corresponding impact on the expected lost sales, where the expectation is taken over the worst-case distribution.

Proposition 4. *Suppose $\boldsymbol{p}^{\boldsymbol{x}*}, Y^{x*}$ be the optimal solution obtained from the conic program to problem (23), where $\boldsymbol{p}_{\boldsymbol{\delta}}^* := \mathbf{E}[\boldsymbol{\delta}^*(\boldsymbol{v}^*)]$; and $Y_\delta^* := \mathbf{E}[\boldsymbol{\delta}^*(\boldsymbol{v}^*)\boldsymbol{v}^{*T}]$. Let $p_{\delta k}^*$ denote the k-element in $\boldsymbol{p}_{\boldsymbol{\delta}}^*$, and $Y_{\delta kk}^*$ the (k,k)-th element of matrix Y_δ^*. Then the decrease in lost sales when T_k^r decreases by a unit is bounded above by $c_k(p_{\delta k}^* - Y_{\delta kk}^*)$.*

As the dual information obtained constitutes an upper bound of the actual impact with unit change in TTR, we could use this information to identify nodes in the supply chain whose reported TTRs will have minimal effect of the performance of lost sales - *the accuracy of the reported TTRs in nodes with low value of $c_k(p_{\delta k}^* - Y_{\delta kk}^*)$ will have minimal impact on the performance of lost sales.*

5.2. Impact of Capacity and Inventory

In a similar vein, we can use the probabilistic interpretation of the worst-case distribution to the conic program to perform sensitivity analysis on other planning parameters. For instance, we can use dual variables to analyze the impact on the change in the vendors' capacities. The optimal dual solution $\boldsymbol{\delta}$, or more specifically, $(T^R - T_k^r(1-v_k))\delta_k$ provides valuable information on the sensitivity of the input c_k. Notice that $T^R - T_k^r(1-v_k)$ is exactly the slack variable s_k^6 defined by constraint (8a). By the same logic as what we have for the sensitivity analysis on vendors' TTR, we expect the effect on vendor's capacity change should depend on the $\mathbf{E}[\boldsymbol{\delta}^*(\boldsymbol{v}^*)\boldsymbol{s}^{\mathbf{6}*\mathsf{T}}(\boldsymbol{v}^*)]$.

Proposition 5. *Let $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}}$ be the submatrix in the optimal solution to the conic program that correspond to $\mathbf{E}[\boldsymbol{\delta}^*(\boldsymbol{v}^*)\boldsymbol{s}^{\mathbf{6}^*}(\boldsymbol{v}^*)]$. Note that $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}}$ is a p-by-p matrix. Let $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}kk}$ be the (k,k)-element in $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}}$, where $k = 1,...,p$. Then the decrease in lost sales when c_k increases by a unit is bounded above by $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}kk}$.*

In this way, we can use the dual information to rule out expanding the capacity of vendor nodes with low $X^*_{\boldsymbol{\delta}\boldsymbol{s}^{\mathbf{6}}kk}$, since the impact on lost sales will be small in these cases. We can apply the same logic to study the sensitivity analysis on inventory level at each plant node.

Proposition 6. *Let $\boldsymbol{p}^*_{\boldsymbol{\beta}}$ be the vector $\boldsymbol{p}^{\boldsymbol{x}*}$ in the optimal solution to the conic program that correspond to $\mathbf{E}[\boldsymbol{\beta}^*(\boldsymbol{v}^*)]$; and $p^*_{\beta i}$ is the i-element in $\boldsymbol{p}^*_{\boldsymbol{\beta}}$. Then we have*

$$\frac{\partial Z^{\hat{e}}(\boldsymbol{r}^*)}{\partial r_i} = -p^*_{\beta i}$$

In this way, we can use the dual information to assess the impact of locally modifying the inventory strategy $\boldsymbol{r}$ by examining the value of $p^*_{\beta_i}$. In the next section, we use numerical simulation to demonstrate that although our results are obtained in the worst-case setting, the insights obtained can still be valuable for the traditional expected lost sales setting, when the risk probabilities are given explicitly.

6. Numerical Studies

We develop the experimental setup using a case study motivated by a large internet service provider serving four different markets. This experiment setup is adapted from a case study of risk analysis in Golany (2014). This company provides four types of internet configurations (CFG1, CFG2, CFG3, and CFG4). In the four markets, to incorporate the issues of flexible supply, we assume that some configurations can be used exchangeably. More specifically, in Market 1, CFG 1 and 4 can be used exchangeably; in Market 2, CFG 1 and 2 can be used exchangeably; in Market 3, CFG 2 and 3 can be used exchangeably; and in Market 4, CFG 3 and 4 can be used exchangeably. Demand rates at the four markets are 0.18, 0.21, 0.2, and 0.21, respectively. Lost sale penalty cost at all customer

Table 2 Bill of Material Information

Item ID	Vendor	Geographic Location	CFG1	CFG2	CFG3	CFG4
ABX	1	US	1	1	1	0
ABN			1	1	1	0
GHY	2	Mexico	1	1	2	1
KIU	3	Mexico	1	0	1	0
PFR	4	US	1	0	1	0
JKI1	5	Mexico	0	3	0	0
JKI2			3	0	0	0
JKI3			2	0	3	2
CFG 1	6	US	1	-	-	-
CFG 2	7	US	-	1	-	-
CFG 3	8	US	-	-	1	-
CFG 4	9	US	-	-	-	1

nodes is 7. The items needed for each configuration, and the corresponding vendor, are listed in Table 2. We map the supply chain network in Figure 5. Supply chain parameters are summarized in Table 3. The disruption probabilities are generated randomly with the sum of probabilities set to 1, so that the average number of vendor nodes disrupted is 1. This is to facilitate comparison with the traditional REI approach, where exactly one vendor node is disrupted in each scenario. Note that in reality, the disruption probabilities should be much lower. We consider also that the disruptions are correlated. We assume Vendor 4 and Vendor 7 are disrupted with correlation coefficient of 0.9; whereas Vendor 1 and Vendor 9 are disrupted with correlation coefficient of 0.9. Instead of simulating $T^R(\boldsymbol{v})$ to obtain the corresponding moments in our experiments, we let the model determine these value in our numerical computation. The machine used to perform all the computations is Dell computer with Intel(R) Core(TM) 3.40 GHz, RAM 8 GB, Microsoft Windows Windows 7 Enterprise cvx Mosek solver. We use Doubly Nonnegative Matrix(DNN) to approximate co-positive matrix in solving the model. The size of the matrix for our numerical study is 141×141 and the computation time for each DNN approximation is around 256s.

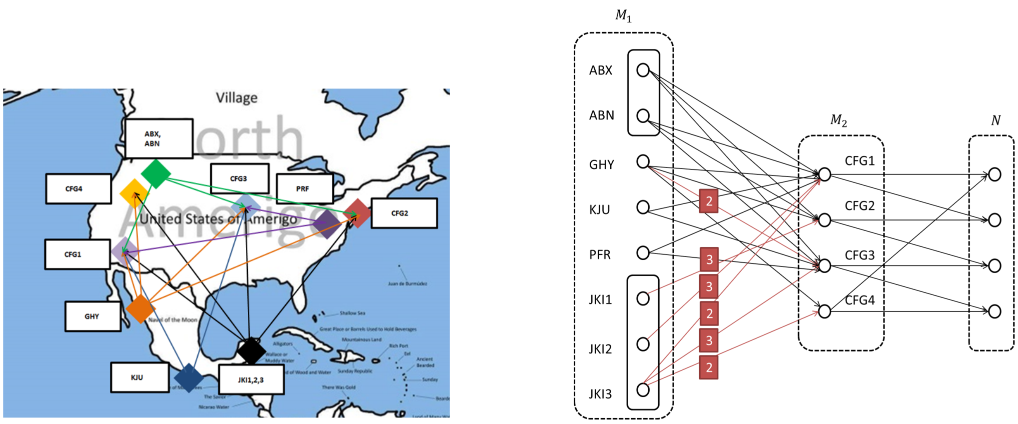


Figure 5 Supply chain network

Table 3 Supply Chain Parameters

Plant	Inventory Cost (100 units)	Vendor	Capacity	Disruption Probability $(1-\mu)$	TTR of vendor
1. ABX	31	1	1.2	0.0667	1
2. ABN	30				
3. GHY	32	2	0.9	0.1333	1.2
4. KJU	29	3	0.5	0.2000	1.4
5. PFR	30	4	0.5	0.1667	1.6
6. JKI1	30	5	2.4	0.0667	1.8
7. JKI2	33				
8. JKI3	31				
9. CFG1	500	6	0.2	0.0667	2
10. CFG2	550	7	0.4	0.1667	2.2
11. CFG3	600	8	0.3	0.0667	2.4
12. CFG4	505	9	0.2	0.0667	2.6

6.1. The Effect of Inventory Budget to Inventory Deployment and WCVaR of Lost Sales

To see how the inventory budget influence the inventory deployment and the corresponding change in WCVaR, we solve Problem (22) again for total inventory budgets ranging from 0.2 to 16, and confidence level $1-\eta$ at 70%, 80% and 90% respectively. The WCVaR of lost sales with respect to

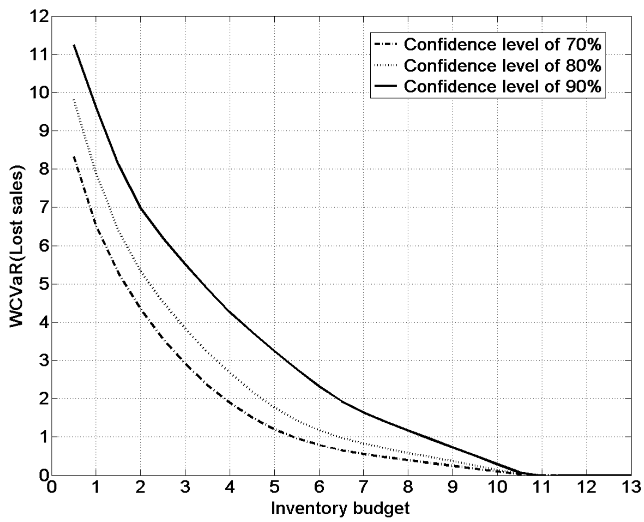


Figure 6 WCVaR of lost sales for different inventory budgets

different level of budget is plotted in Figure 6.

Note that it is easy to reduce lost sales to zero with a high enough inventory budget. In fact, with the maximum TTR at 2.6, we can pre-position the total demand for the 4 markets for 2.6 unit time at the 4 plant nodes in the set M_2, incurring a total inventory cost of

$$2.6 \times (0.18 \times 5 + 0.21 \times 5.5 + 0.2 \times 6 + 0.21 \times 0.505) = 11.203.$$

The challenge is therefore to pre-position the inventory with a much smaller budget. Interestingly, our WCVaR model essentially recovers this insight, and shows the diminishing return on inventory budget on lost sales mitigation - when total inventory budget is small, a small increase in budget can significantly decrease the WCVaR of lost sales. However, when the inventory budget reaches 10.5 to 11, the lost sales effectively reduces to 0, and any further increase in inventory budget will not affect the WCVaR of lost sales. There is thus a diminishing returns to the value of additional inventory budget in controlling for WCVaR.

6.2. Inventory Strategy Comparison with REI Model

We compare the optimal inventory deployment levels under two environments, one without disruption distribution information (LP based on the REI model, ignoring the disruption risk estimates), and the other one with limited distribution information (COP model).

For the first case, optimal inventory allocations and total inventory budgets are obtained by solving Problem (11). The total inventory budget obtained is 2.6186, with inventory allocation as shown by the white bar in Figure 7. Note that this strategy is obtained with the optimistic assumption that there is a disruption to at one vendor node in the system. This is often justified because of the low disruption risk to each vendor node in the system. We use this case as a benchmark to evaluate the performance of the COP model for this experimental setup.

With this fixed inventory budget (2.6186), and given the first-two moment information, how would the optimal inventory allocation change using the COP model? We solve the fixed inventory budget problem (22) with budget of $br = 2.6186$. To study the risk-aversion effect, we test our model under three cases of confidence levels, 70% ($\eta = 0.3$), 80% ($\eta = 0.2$) and 90% ($\eta = 0.1$). The new optimal inventory allocation strategies are compared with the one obtained from LP model in Figure 7. We further analyze how large the lost sales perform under inventory strategies obtained by the optimistic-LP-based model using REI, and the pessimistic-based-COP model assuming worst-case distributions. Specifically, given the disruption moment information, we simulate the performance of different inventory strategies and the accompanied lost sales. When we assume that both models have the same inventory budget of 2.6186, the cumulative distribution functions of lost sales under these inventory strategies are given in Figure 8 with average lost sales, standard deviations, 70% CVaR, 80% CVaR, and 90% CVaR given in Table 4. From both the CDFs of

Table 4 Statistics of simulated lost sales under the same budget

	Mean	STD	70% CVaR	80% CVaR	90% CVaR
COP Model (Confidence level 70%)	0.5633	1.1475	1.8720	2.4019	3.5173
COP Model (Confidence level 80%)	0.5372	1.1286	1.7846	2.3869	3.4562
COP Model (Confidence level 90%)	0.5612	1.1621	1.8623	2.4642	3.6396
LP Model	0.9569	1.9128	3.1897	4.6952	5.1541

lost sales and mean-variance comparison, we can see that the COP models perform significantly

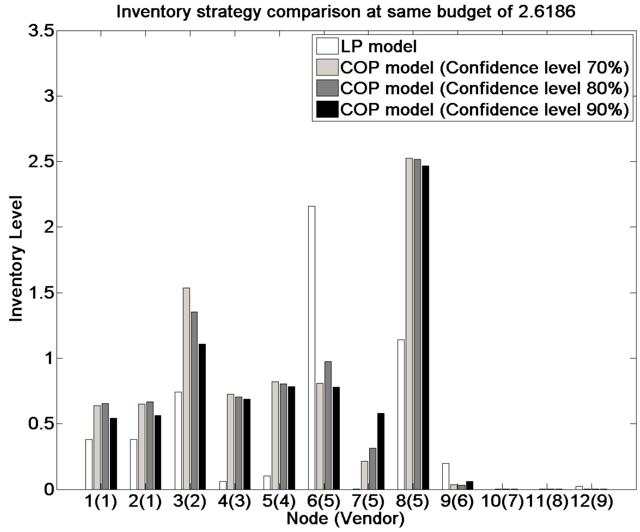


Figure 7 Inventory levels for each plant under same budget

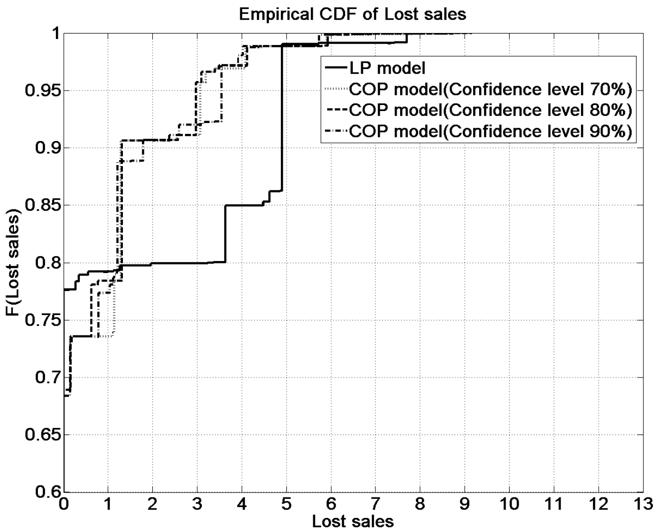


Figure 8 CDFs of simulated lost sales using inventory obtained from two models under the same budget

better than LP model. The optimistic assumption that at most one vendor node can be disrupted can therefore lead to poor inventory deployment strategy. Our numerical example shows that the incorporation of risk estimates (probabilities of disruption) can therefore be valuable for this problem, even if the model can only be solved for the worst-case setting.

6.3. Inventory Strategy Comparison with the Stochastic Programming Model

Based on the disruption probabilities at each node, we first generate random scenarios with up to 4 vendors being disrupted, assuming (wrongly) that the disruption risk of each vendor node in the supply chain are independent. We remove all scenarios with more than four vendors being disrupted. Note that the probability that five or more vendors are simultaneously disrupted is less than 0.004% according to the disruption probabilities in Table 3. In this way, we completely enumerate 255 scenarios to build the stochastic programming model in (11). We then apply Benders decomposition method to find the optimal solution.

In order to make a fair comparison, we compare the inventory strategies obtained from stochastic model to the one obtained from COP model under the belief that disruptions are uncorrelated. This is to ensure both models use the same amount of information on the disruption risk assessments.

The supply chain network and all the problem parameters are unchanged. We set the total inventory budget to be 2.6186. We obtain three sets of inventory strategies with CVaR confidence level of 70%, 80%, and 90%. The true underlying disruption distributions are however correlated, as in our base case. Specifically, both vendor nodes 4 and 7, and vendor nodes 1 and 9, are correlated, each with correlation coefficient of 0.9.

Note that we do not assume the true disruption probabilities among companies are independent. This is because if so, the stochastic programming model, which is solved assuming (correctly) disruptions are independent, will always outperform the worst-case COP model under simulation, since the stochastic programming approach solves the right optimization problem.

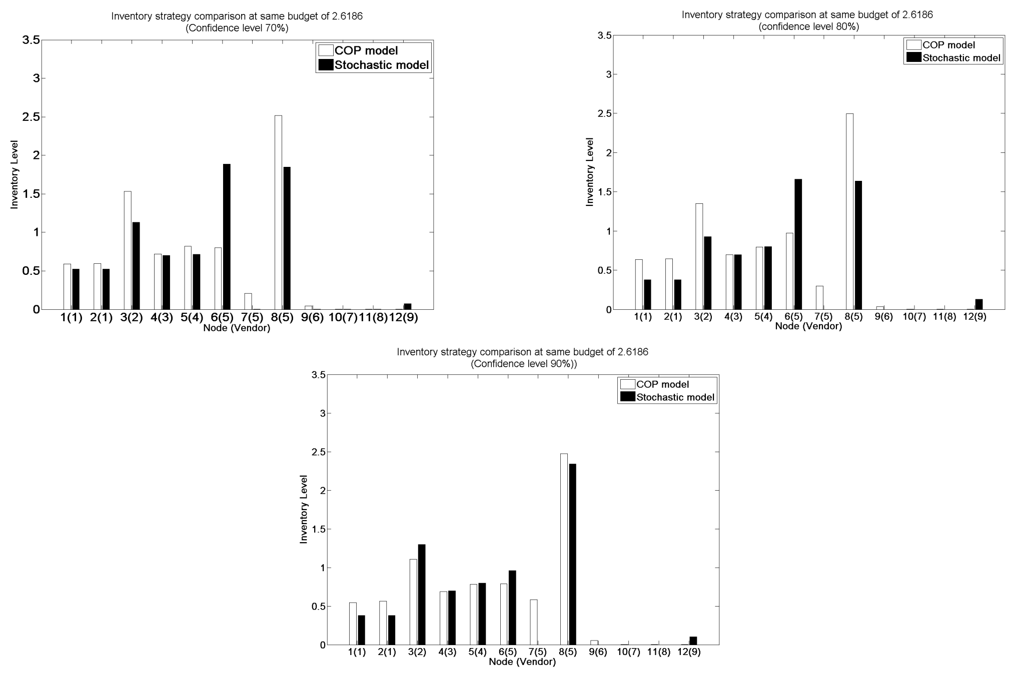


Figure 9 Inventory Strategy Comparison

In Figure 9, we present the inventory deployment strategy obtained from the stochastic programming model and the one obtained from COP models. To compare the performance of these two strategies, we conduct simulation study by assuming the true disruption distribution are correlated. We plot the cumulative distribution functions for different inventory strategies and under different confidence levels in Figure 10 with average lost sales, standard deviations, 70% CVaR, 80% CVaR, and 90% CVaR in Table 5.

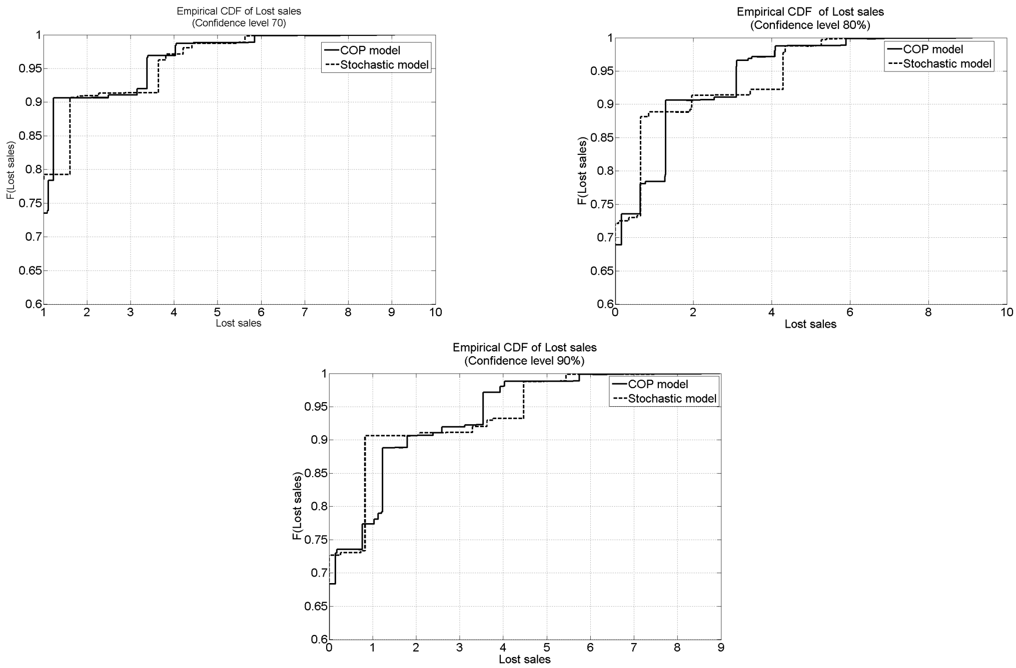


Figure 10 Simulated Lost Sales CDFs Comparison

Table 5 Statistics of simulated lost sales under the same budget

	Mean	STD	70% CVaR	80% CVaR	90% CVaR
COP Model (Confidence level 70%)	0.5673	1.1675	1.8851	2.4305	3.6178
Stochastic Model(Confidence level 70%)	0.5623	1.2343	1.8745	2.6999	3.7901
COP Model (Confidence level 80%)	0.5383	1.1339	1.7884	2.3937	3.4969
Stochastic Model(Confidence level 80%)	0.5341	1.2635	1.7803	2.4339	4.0535
COP Model (Confidence level 90%)	0.5605	1.1626	1.8603	2.4669	3.6391
Stochastic Model(Confidence level 90%)	0.5476	1.2681	1.8254	2.4472	4.0676

From the CDF plots, we can see the CDF curves corresponding to the inventory strategies from the stochastic programming model are more skewed compared with the ones corresponding to the COP model. This implies the COP model performs slightly better than the stochastic model in the tail. In Table 5, we see that the COP model is inferior to the stochastic programming model in terms of mean lost sales. However, in terms of spread (standard deviation), COP model outperforms the stochastic programming model. At the same time, COP model controls the tails much better than the stochastic model (c.f. CVaR statistics). Since downside risk is more critical than simply mean lost sales, COP model is better in controlling for the disruption risk.

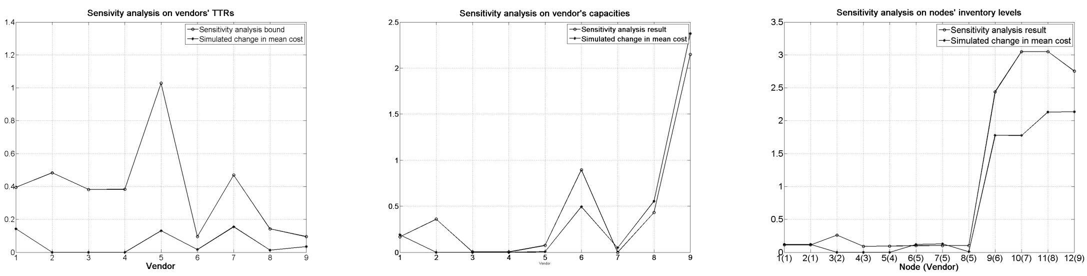
In terms of computation time, COP model outperforms the stochastic model significantly. The computation time to obtain the inventory strategy from stochastic model is about 45mins, compared with less than 5 mins in the case of running the COP model.

6.4. Sensitivity Analysis

In this section, given the optimal inventory deployment with the fixed inventory budget 2.6186 obtained by the COP model, we would like to analyze the effects of the key planning parameters on the performance of the supply chain. The inventory strategy we use here is the one obtained with confidence level 70%. By setting θ to be 0, we first solve Problem (23) to obtain the corresponding primal and dual optimal solution, to perform the sensitivity analysis on vendors' TTRs, capacities and inventory deployment.

To validate this result, we use simulation to check the impact when these parameters change one at a time. For instance, for the analysis on TTR, we have 9 cases where each case corresponds to decreasing one vendor's TTR by 0.1 units. We sample 10^5 disruption scenarios to estimate the mean lost sales for each of these cases, and obtain the estimated change in performance level (lost sales). We perform the same analysis for the cases when capacities or inventory positions are changed, by increasing the vendor's capacity by 0.1 units,and inventory position by 0.001 units each.

The comparisons are shown in Figure 11. The simulated changes in the mean lost sales are normalized to be the change with respect to a unit change of TTR, capacity or inventory position.



(a) Sensitivity analysis on vendors' TTRs

(b) Sensitivity analysis on vendors' capacities

(c) Sensitivity analysis on nodes' inventory levels

Figure 11 Sensitivity analysis

It is interesting that while vendors 6-9 are those with the highest TTR, our sensitivity analysis on TTR (cf. Figure 11a) shows that the effect of decreasing the TTR at Vendor node 6, 8 and 9 have the lowest impact on the performance of the supply chain, since our analysis yields the lowest upper bound for these vendor nodes. Hence the TTR from these nodes, despite having the largest value, are not critical to the performance of the system under disruption. In fact, the simulation results confirmed this finding, since changing the TTR at these nodes has a negligible impact on the expected performance on lost sales. On the other hand, the upper bound on the effect of the TTR of Vendor 5 is shown to be large in the supply chain. This is not surprising, since Vendor 5 is responsible for 3 items in the set M_1, and holds a large amount of the strategic inventory positioned in the system.

More interestingly, in the case of capacity parameters, the situation is reversed and our sensitivity analysis shows that Vendor 6, 8 and 9 now played the most important role in the performance of the supply chain, with the simulation results confirming this findings (cf. Figure 11b). The capacity parameters for other vendors do not appear to be significant, except possibly Vendor 1 and 2, based on sensitivity analysis. Simulation confirms that capacity parameter at Vendor 1, but not Vendor 2, has a non-negligible impact on the performance.

The situation with the impact of inventory positioning is more intuitive - vendors 6, 7, 8 and 9 are significant, and a slight increase in the inventory positions in these nodes will have a larger impact of the performance in lost sales. This is arguably due to the fact that these nodes are closer to markets, and the inventory will not be destroyed in the case of disruption, unlike the case with capacity. This possibly explains why Vendor 7 is now a significant node in the supply chain, despite having also a larger disruption probability. The inventory positioning at all other nodes are not as important, according to our sensitivity analysis, and again confirmed by simulation.

6.5. Effect of Budget: Optimal Inventory Strategy is not monotone

Given that the budget for strategic inventory positioning may change from time to time, it is important to understand how we could build up the strategic inventory position over time. To understand this effect, we obtain the optimal inventory positioning strategy for a range of budgets. We show the different inventory positions for different levels of inventory budget in Figure 12, with confidence level 70%. Note that our earlier numerical results suggest that the inventory positions are not sensitivity to the confidence level, since the solutions are similar even at confidence levels 80% and 90%.

We can see tier-two vendors (Plant Node 1-8) and tier-one vendors (Plant Node 9-12) respond to the increase in total budget in different manner. When the total inventory budget is sufficiently low, it is optimal to spread any additional inventory investment across all vendors to increase their respective inventory level. However, when the total amount of budget crosses a certain threshold (in our case, it is about 4 to 6), it is optimal to continue to invest more in the inventory levels of tier-one vendors, and also to strategically reduce the inventory positions in selected tier-two vendors. This follows from the complex trade-offs between the higher inventory holding cost of tier-one vendors, and their relative proximity to end demands compared to tier-2 vendors. Essentially the optimal inventory positioning strategy uses the following principle - when the total inventory budget is small, the supply chain has to invest more on the tier-two vendors. When the budget is sufficient, the balance will be titled towards the tier-one vendors, allowing the supply chain to invest more on holding inventories for these vendors. The numerical results also reveal a consistent

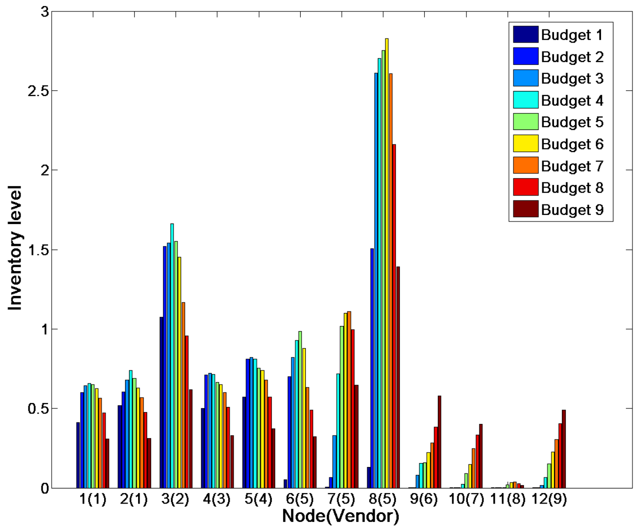


Figure 12 Inventory levels under different inventory budgets

pattern - over the range of budgets considered, Plant Node 8 and 3 hold the largest and second largest share of the amount of strategic inventory available, with close to zero strategic inventory maintained at Plant Node 11.

7. Concluding Remarks

In this paper, we introduce a risk mitigation framework, which incorporates disruption risk estimates into supply chain planning. Our mitigation framework is developed based on the supply chain resilience curve on the performance of lost sales. We present a framework to determine the optimal inventory allocation strategy such that the anticipated lost sales are minimized. Specifically, we assume the first-two moment of the disruption distribution are known, and by adopting the notion of distributionally robust model, we obtained the optimal inventory allocation across the supply chain which gives minimal value of WCVaR of the lost sales.

We showed that this problem can be fully characterized by a co-positive program, which can be solved via SDP relaxation. Moreover, our distributionally robust model can be used to perform sensitivity analysis, and to estimate the changes in the worst-case expected lost sales when certain supply chain parameters change. We finally apply our framework to a numerical study. We show that in both cases when the disruptions are either independent or correlated, the optimal inventory strategy obtained by our co-positive program model outperforms that obtained from the traditional REI model.

There are several interesting implications our numerical study reveals. Specifically, our numerical results highlight the important role that total inventory budget plays in supply chain risk management. On one hand, we show there is a diminishing return effect to the value of additional inventory budget in controlling WCVaR of lost sales. On the other hand, we find that the optimal inventory strategy is not always monotone with respect to total inventory budget. The single-crossing condition does not hold in this case. When the total inventory budget passes certain thresholds, the optimal inventory for certain nodes may begin to drop.

We also would like to make some remarks and highlight several limitations in our model. Firstly, our risk mitigation model does not incorporate the factor of lead times. Introducing lead times, even deterministic lead times, would impose another layer of difficulties in modeling and solutions. The reason why we find it is not essential to include lead times is that the inventory allocation strategy we proposed in our model is specifically used to hedge against random supply shortage. It can be viewed as an extra buffer held in addition to those inventory already installed in the supply chain system, which is used to dampen the variability in demand and lead times. We decouple the two sources of uncertainties and only focus on determining the inventory position after receiving risk alerts on possible disruption threats.

Secondly, we solve the SDP relaxation of the corresponding co-positive and completely positive programs using the standard Doubly-Non-Negative (DNN) relaxation. In general, the gap between the exact and the relaxed model may be large, and we have introduced additional constraints that can be added to the formulation to reduce the gap and ensure convergence. Note that there are cases in which exact SDP formulation can be found to be equivalent to the co-positive/completely positive program. Specifically, in Natarajan and Teo (2016), they explicitly showed that the order-statistic problem is one such case. It is not difficult to show that when the supply network is a balanced serial chain, our risk management problem is an order statistic problem. Hence, we can have an exact SDP model for this special structure. Whether there any other networks possessing such tight formulations or, more generally, whether there are any conditions under which we can have a tight formulation, is a promising future research question.

Finally, in terms of numerical study, CVX Mosek cannot be used to solve large scale DNN problems. When the supply chain is large, one possible way to apply our framework is to decompose the network into blocks, and solve each block one at a time to fix the inventory levels. It will be interesting to see if there is a more compact way to model the supply chain disruption problem.

Endnotes

1. In fact, we can extend the result to the case when the support set is unbounded by explicitly characterizing the recession cone of the completely positive cone. For simplicity, in our paper, we only study the case when the support set is bounded.

References

[1] Apparel (March 16, 2015). Supply Chain Disruption Tops Manufacturers' Concerns, BSI Finds. Retrieved August 15, 2015, http://apparel.edgl.com/news/Supply-Chain-Disruption-Tops-Manufacturers–Concerns,-BSI-Finds99154.

[2] Arreola-Risa, A. and DeCroix, G.A. (1998) Inventory management under random supply disruptions and partial backorders. *Naval Research Logistics*, **45**(7): 687-703.

[3] Berman, A., Shaked-Monderer, N. (2003) Completely Positive Matrices. *World Scientific, Singapore.*

[4] Bertsimas, D., Sethuraman, J.(2000) Moment problems and semidefinite optimization. In *Handbook of semidefinite programming* 469-509. Springer US.

[5] Braunscheidel, M.J., Suresh, N.C. (2009) The organizational antecedents of a firm's supply chain agility for risk mitigatoin and response. *Journal of Operations Management*, **27**(2): 119-140.

[6] Burer, S. (2009) On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, **120**(2): 479-495.

[7] Craighead, C.W., Blackhurst, J., Rungtusanatham, M.J., Handfield, R.B. (2007) The severity of supply chain disruptions: design characteristics and mitigation capabilities. *Decision Sciences*, **38**(1): 131-156.

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

[9] Ferrer, J., Karlberg, J., Hintlian, J.(2007) Integration: The key to global success. *Supply Chain Management Review*, **11**(2).

[10] Geraint, J. (2014) Innovative approaches to supply chain risk. *SCM World Research Report.*

[11] Golany, S.Y. (2011) Enhancing service providers reliability by mitigating supply chain risk: The case of telecommunication networks. *Master Thesis MIT.*

[12] Gurnani, H., Mehrota, A., Ray, S. (2012) Supply chain disruptions: Theory and practice of managing risk. *Springer-Verlag London.*

[13] Kleindorfer, P.R., Saad, G.H. (2005) Managing disruption risks in supply chains. *Production and Operations Management*, **14**(1): 53-68.

[14] Hanasusanto, G.A., Kuhn, D. (2016) Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls. *Working Paper.*

[15] Hopp W., Yin, Z. (2006) Protecting supply chain networks against catastrophic failures. *Working Paper.*

[16] de Klerk, E., Pasechnik, D.V. (2002) Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, **12**(4): 875-892.

[17] Levermann A. (2014) Climate economics: Make supply chains climate-smart. *Nature*, **506**, 27-29.

[18] Meyer, R. R., Rothkoft, M. H., Smith, S. A. (1979) Reliability and inventory in a production-storage system. *Management Science*, **25**(8): 799-807.

[19] Natarajan, K., Teo, C. P. (2017) On reduced semidefinite programs for second order moment bounds with applications. *Mathematical Programming*, **161**(1-2): 487-518.

[20] Natarajan, K., Teo, C. P., Zheng, Z. (2011) Mixed zero-one linear programs under objective uncertainty: a completely positive representation. *Operations Research*, **59**(3): 713-728.

[21] Parrilo, P. (2000) Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. *Ph.D. Dissertation* California Institute of Technology.

[22] Pflug, G.C. (2000) Some remarks on the value-at-risk and the conditional value-at-risk. *Probabilistic constrained optimization*, 272-281.

[23] Rockafellar R. T., Uryasev S. (2000) Optimization of conditional Value-at-Risk. *Journal of Risk* **2**:21-42.

[24] Rockafellar R. T., Uryasev S. (2002) Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance* **26**:1443-1471

[25] Sheffi, Y. (2015). Preparing for disruptions through early detection. *MIT Sloan Management Review*, **57**(1): 31-42.

[26] Simchi-Levi, D., Schmidt, W., Wei, Y. (2014) From superstorms to factory fires: Mangaing unpredictable supply chain disruptions. *Harvard Business Review*, **92**(1-2): 96-101.

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

[28] Simchi-Levi, D., He Wang, Wei, Y. (2015b) Increasing supply chain robustness through process flexibility and inventory. *Working Paper.*

[29] Simchi-Levi, D. (2015c) Find the weak link in your supply chain. *Harvard Business Review.*

[30] Song, J.S., Zipkin, P. H. (1996) Inventory control with information about supply conditions. *Management Science*, **42**(10): 1409-1419.

[31] Tang, C. (2006) Robust strategies for mitigating supply chain disruptions. *International Journal of Logistics*, **9**(1): 33-45.

[32] Tomlim, B., Snyder L. (2007). On the value of a threat advisory system for managing supply chain disruptions. *Working Paper.*

[33] Yan, T., Choi, T.Y., Kim, Y., Yang. Y., (2015) A theory of the nexus supplier: a critical supplier from a network perspective, *Journal of Supply Chain Management*, **51**(1): 52-66

[34] Zhu, S., Fukushima, M. (2009) Worst-case conditional value-at-risk with application to robust portfolio management. *Operations research*, **57**(5): 1155-1168.

Proofs of Main Results

EC.1. Proofs of Main Results

Proof of Proposition 1 *Necessity:* Let $(\boldsymbol{w}, X^w)$ be a element in the set $\mathcal{M}$. Consider the decomposition of the completely positive matrix:

$$\begin{pmatrix} 1 & \boldsymbol{w}^\top \\ \boldsymbol{w} & X^w \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \zeta_k \\ \hat{\boldsymbol{v}}_k \end{pmatrix} \begin{pmatrix} \zeta_k \\ \hat{\boldsymbol{v}}_k \end{pmatrix}^\top$$

where $\zeta_k \in \mathcal{R}_+$, $\sum_{k\in\kappa} \zeta_k^2 = 1$, $\hat{\boldsymbol{v}}_{\boldsymbol{k}} \in \mathcal{R}_+^p, \forall k \in \kappa$. Let $\kappa = \kappa_+ \cup \kappa_0$, where $\kappa_+ = \{k \in \kappa \mid \zeta_l > 0\}$, and $\kappa_0 = \{k \in \kappa \mid \zeta_l = 0\}$. We have $\boldsymbol{w} = \sum_{k\in\kappa} \zeta_k \hat{\boldsymbol{v}}_{\boldsymbol{k}}$, and $X^w = \sum_{k\in\kappa} \hat{\boldsymbol{v}}_k \hat{\boldsymbol{v}}_k^\top$.

Following the result in Proposition 3.1 and Proposition 3.2 in Natarajan et al. (2011) (also from Burer(2009)), we have $\forall k \in \kappa_+, \frac{\hat{\boldsymbol{v}}_k}{\zeta_k}$ is feasible to linear constraints and binary constraints in support set $\mathcal{D}$; and $\hat{\boldsymbol{v}}_k = 0, \forall k \in \kappa_0$.

Before we proceed with the proof, we would like to first establish the following lemma.

Lemma EC.1. *$\forall k \in \kappa_+, \frac{\hat{\boldsymbol{v}}_k}{\zeta_k}$ is feasible to the quadratic constraint, $(M_2\tilde{\boldsymbol{v}}) \circ (M_3\tilde{\boldsymbol{v}}) = 0$, in support set $\mathcal{D}$,*

Proof. $\text{diag}(M_2 X^w M_3^\top) = 0$ can be rewritten as $\sum_{k\in\kappa_+} \zeta_k^2 (M_2 \frac{\hat{\boldsymbol{v}}_k}{\zeta_k})(M_3 \frac{\hat{\boldsymbol{v}}_k}{\zeta_k})^\top = 0$. Because M_2 and M_3 satisfies the condition that for all $\boldsymbol{v}$ feasible to linear constraint $M_1\tilde{\boldsymbol{v}} = \boldsymbol{b}_1^v$, $M_2\boldsymbol{v} \geq 0$ and $M_3\boldsymbol{v} \geq 0$. We have $M_2\frac{\hat{\boldsymbol{v}}_k}{\zeta_k} \geq 0$, and $M_3\frac{\hat{\boldsymbol{v}}_k}{\zeta_k} \geq 0, \forall k \in \kappa_+$. Therefore, the constraint $\text{diag}(M_2 X^w M_3^\top) = 0$ will force $(M_2\frac{\hat{\boldsymbol{v}}_k}{\zeta_k})(M_3\frac{\hat{\boldsymbol{v}}_k}{\zeta_k})^\top = 0, \forall k \in \kappa_+$. Q.E.D.

With this lemma, we conclude $\forall k \in \kappa_+, \frac{\hat{\boldsymbol{v}}_k}{\zeta_k}$ is feasible to all constraints in support set $\mathcal{D}$. We can rewrite the decomposition as

$$\begin{pmatrix} 1 & \boldsymbol{w}^\top \\ \boldsymbol{w} & X^w \end{pmatrix} = \sum_{k\in\kappa_+} \begin{pmatrix} 1 \\ \frac{\hat{\boldsymbol{v}}_k}{\zeta_k} \end{pmatrix} \begin{pmatrix} 1 \\ \frac{\hat{\boldsymbol{v}}_k}{\zeta_k} \end{pmatrix}^\top$$

We can construct a distribution of a binary random vector $\tilde{\boldsymbol{v}}$ in the following way. $P(\tilde{\boldsymbol{v}} = \frac{\hat{\boldsymbol{v}}_k}{\zeta_k}) = \zeta_k^2, \forall k \in \kappa_+$. The probability of $\tilde{\boldsymbol{v}}$ taking values other than $\frac{\hat{\boldsymbol{v}}_k}{\zeta_k}, \forall k \in \kappa_+$ is 0. This binary random vector satisfies all the constraints in support set $\mathcal{D}$ almost surely. We then have $\boldsymbol{\mu} = \boldsymbol{w} = \sum_{k\in\kappa_+} P(\tilde{\boldsymbol{v}} = \frac{\hat{\boldsymbol{v}}_k}{\zeta_k})\frac{\hat{\boldsymbol{v}}_k}{\zeta_k}; \Sigma = X^w = P(\tilde{\boldsymbol{v}} = \frac{\hat{\boldsymbol{v}}_k}{\zeta_k})\frac{\hat{\boldsymbol{v}}_k}{\zeta_k}\frac{\hat{\boldsymbol{v}}_k}{\zeta_k}^\top$. Hence, we can see $\boldsymbol{\mu}$ and Σ are feasible first-two moments of this binary random vector $\tilde{\boldsymbol{v}}$.

Sufficiency: Supposing $(\boldsymbol{\mu}, \Sigma)$ is a feasible $(n, 2, \mathcal{D})$-moment sequence, we have $\boldsymbol{\mu} = \mathbf{E}[\tilde{\boldsymbol{v}}]$; $\Sigma = \mathbf{E}[\tilde{\boldsymbol{v}}\tilde{\boldsymbol{v}}^\top]$. Define $\boldsymbol{w} = \mathbf{E}[\tilde{\boldsymbol{v}}]$, $X^w = \mathbf{E}[\tilde{\boldsymbol{v}}]$. Then $\boldsymbol{w} = \boldsymbol{\mu}^v; X^w = \Sigma^v$. By taking expectation of all the constraints in $\mathcal{D}$, we have $M_1\boldsymbol{w} = \boldsymbol{b}_1^v$, $\text{diag}(M_1 X^w M_1^\top) = \boldsymbol{b}_1^v \cdot \boldsymbol{b}_1^v$, and $\text{diag}(M_2 X^w M_3^\top) = 0$. Due to the fact that $\tilde{v}_i \in \{0,1\}$ we have $\tilde{v}_i = \tilde{v}_i^2$, Taking expectations gives $w_i = X^w_{ii}, \forall i \in \mathcal{B}^v$. Finally, $\tilde{\boldsymbol{v}} \geq \mathbf{0}$ almost surely implies $\begin{pmatrix} 1 & \boldsymbol{w}^\top \\ \boldsymbol{w} & X^w \end{pmatrix} \succcurlyeq_{cp} 0$ This completes the proof. Q.E.D.

Proof of Theorem 1. By the construction the Problem (4), it is a relaxation of Problem (2). To see the equivalence, let $\boldsymbol{p}^{x*}, \boldsymbol{w}^*, X^{x*}, X^{w*}$, and Y^{x*} be the optimal solution to Problem (4), and consider the rank-1 decomposition of the completely positive matrix at this optimal solution, i.e.,

$$\begin{pmatrix} 1 & \boldsymbol{p}^{x*\top} & \boldsymbol{w}^{*\top} \\ \boldsymbol{p}^{x*} & X^{x*} & Y^{x*} \\ \boldsymbol{w}^* & {Y^{x*}}^\top & X^{w*} \end{pmatrix} = \sum_{k\in\kappa} \begin{pmatrix} \alpha_k \\ \boldsymbol{\beta_k} \\ \boldsymbol{\gamma_k} \end{pmatrix} \begin{pmatrix} \alpha_k \\ \boldsymbol{\beta_k} \\ \boldsymbol{\gamma_k} \end{pmatrix}^\top = \sum_{k\in\kappa} \alpha_k^2 \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta_k}}{\alpha_k} \\ \frac{\boldsymbol{\gamma_k}}{\alpha_k} \end{pmatrix} \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta_k}}{\alpha_k} \\ \frac{\boldsymbol{\gamma_k}}{\alpha_k} \end{pmatrix}^\top$$

where $\alpha_k \in \mathcal{R}_+$, $\sum_{k\in\kappa} \alpha_k^2 = 1$, $\boldsymbol{\beta_k}, \boldsymbol{\gamma_k} \in \mathcal{R}^n_+, \forall k \in \kappa$. Similarly, let $\kappa = \kappa_+ \cup \kappa_0$, where $\kappa_+ = \{k \in \kappa \mid \alpha_k > 0\}$, and $\kappa_0 = \{k \in \kappa \mid \alpha_k = 0\}$.

From the result in Proposition 3.1, Proposition 3.2 in Natarajan et al.(2011) and Lemma EC.1, we have $\frac{\boldsymbol{\beta_k}}{\alpha_k}, \forall k \in \kappa_+$ be the feasible solutions to Problem (2) and $\boldsymbol{\beta_k} = \mathbf{0}, \forall k \in \kappa_0$. Similarly, these three lemmas also implies $\frac{\boldsymbol{\gamma_k}}{\alpha_k}, \forall k \in \kappa_+$ be the feasible solutions to the feasible moment problem and $\boldsymbol{\gamma_k} = \mathbf{0}, \forall k \in \kappa_0$. We can rewrite the decomposition as

$$\begin{pmatrix} 1 & \boldsymbol{p}^{x*\top} & \boldsymbol{w}^{*\top} \\ \boldsymbol{p}^{x*} & X^{x*} & Y^{x*} \\ \boldsymbol{w}^* & {Y^{x*}}^\top & X^{w*} \end{pmatrix} = \sum_{k\in\kappa_+} \alpha_k^2 \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta_k}}{\alpha_k} \\ \frac{\boldsymbol{\gamma_k}}{\alpha_k} \end{pmatrix} \begin{pmatrix} 1 \\ \frac{\boldsymbol{\beta_k}}{\alpha_k} \\ \frac{\boldsymbol{\gamma_k}}{\alpha_k} \end{pmatrix}^\top$$

Assume a random vector $\boldsymbol{v}^*$ and its corresponding feasible solutions $\boldsymbol{x}^*(\boldsymbol{v}^*)$ follow the joint distribution as below.

$$P((\boldsymbol{x}^*(\tilde{\boldsymbol{v}}^*), \tilde{\boldsymbol{v}}^*) = (\frac{\boldsymbol{\beta}_k}{\alpha_k}, \frac{\gamma_k}{\zeta_k})) = \alpha_k^2, \forall k \in \kappa_+.$$

It can be easily verified this is a valid and feasible distribution which satisfies the first-two moments of $\tilde{\boldsymbol{v}}$. By the same argument as in Natarajan et al.(2011), Problem (2) is equivalent to Problem (4). Q.E.D.

Proof of Proposition 2. We establish strong duality via constructing an interior point of the following co-positive cone in Problem (5). We first set $\boldsymbol{\nu}_v = \mathbf{0}, \boldsymbol{\psi}_v = \mathbf{0}, \boldsymbol{\phi}_v = \mathbf{0}, \boldsymbol{\phi}_{xv} = \mathbf{0}, \boldsymbol{\psi}_x = \mathbf{0}, \boldsymbol{\lambda} = \mathbf{0}, \boldsymbol{\epsilon}_{xv} = \mathbf{0}, \boldsymbol{\epsilon}_v = \rho\mathbf{1}, \Theta_v = \rho\mathbb{I}$, and $\boldsymbol{\epsilon}_x = \mathbf{1}$, where $\mathbb{I}$ is the identity matrix. The co-positive matrix becomes

$$CO_0 = \begin{pmatrix} \rho & \mathbf{0}^\top & \frac{1}{2}({A_1}^\top \boldsymbol{\phi}_x - \boldsymbol{c}_1)^\top \\ \mathbf{0} & \rho(\mathbb{I} + M_1^\top M_1) & -\frac{C_2}{2}^\top \\ \frac{1}{2}({A_1}^\top \boldsymbol{\phi}_x - \boldsymbol{c}_1) & -\frac{C_2}{2} & {A_1}^\top A_1 - C_3 \end{pmatrix}$$

We then make use of the co-positive Schur complement result in Hanasusanto and Kuhn (2017).

Lemma EC.2 **(Co-positive Schur Complement (Hanasusanto and Kuhn (2017)))**. *Consider a symmetric matrix*

$$D = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}$$

with $A \succ 0$. Then $D \succ_{co} 0$ if $C - B^\top A^{-1} B \succ_{co} 0$.

According to Schur complement, because $\mathbb{I}+M_1^\top M_1 \succ 0$, we have $CO_0 \succ_{co} 0$ if

$$A_1{}^\top A_1 - C_3 \succ_{co} \frac{1}{\rho}\begin{pmatrix} \frac{1}{2}(A_1{}^\top \boldsymbol{\phi}_x - \boldsymbol{c}_1)^\top \\ -\frac{C_2}{2}{}^\top \end{pmatrix}^\top \begin{pmatrix} 1 & \boldsymbol{0}^\top \\ \boldsymbol{0} & \mathbb{I}+M_1^\top M_1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{2}(A_1{}^\top \boldsymbol{\phi}_x - \boldsymbol{c}_1)^\top \\ -\frac{C_2}{2}{}^\top \end{pmatrix}.$$

It holds when we set ρ to be a very large number under the condition that $A_1{}^\top A_1 - C_3 \succ_{co} 0$. Q.E.D.

Proof of Proposition 3. When there exists a partition of decision variable $\boldsymbol{x}$ into two subvectors $\begin{pmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{pmatrix}$ such that the objective of Problem (1) can be rewritten as $\hat{\boldsymbol{c}}_1^\top \boldsymbol{x}_1 + \boldsymbol{v}^\top \hat{C}_2 \boldsymbol{x}_1 + \boldsymbol{x}_2^\top \hat{C}_3 \boldsymbol{x}_1$, we rewrite the completely positive program which have decision variables corresponding to decomposed subvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ as follows.

$$
\begin{array}{lll}
max\ \hat{\boldsymbol{c}}_1{}^\top \boldsymbol{p}^1 + \hat{C}_2 \cdot Y^1 + \hat{C}_3 \cdot X^{12} & & \\
s.t.\ \underline{\textit{Constraints on Decision Variables}} & & \underline{\textit{Dual Variables}} \\
\begin{pmatrix} A_1^1 & O \end{pmatrix}\begin{pmatrix}\boldsymbol{p}^1\\ \boldsymbol{p}^2\end{pmatrix} = \boldsymbol{b}_1^{(1)} & & \boldsymbol{\phi}_x^{(1)} \\
\begin{pmatrix} O & A_1^2 \end{pmatrix}\begin{pmatrix}\boldsymbol{p}^1\\ \boldsymbol{p}^2\end{pmatrix} = \boldsymbol{b}_1^{(2)} & & \boldsymbol{\phi}_x^{(2)} \\
\mathrm{diag}\big(\begin{pmatrix} A_1^1 & O \end{pmatrix}\begin{pmatrix} X^1 & X^{12} \\ X^{12\top} & X^2 \end{pmatrix}\begin{pmatrix} A_1^1 & O \end{pmatrix}^\top\big) = \boldsymbol{b}_1^{(1)} \circ \boldsymbol{b}_1^{(1)} & & \boldsymbol{\epsilon}_x^{(1)} \\
\mathrm{diag}\big(\begin{pmatrix} O & A_1^2 \end{pmatrix}\begin{pmatrix} X^1 & X^{12} \\ X^{12\top} & X^2 \end{pmatrix}\begin{pmatrix} O & A_1^2 \end{pmatrix}^\top\big) = \boldsymbol{b}_1^{(2)} \circ \boldsymbol{b}_1^{(2)} & & \boldsymbol{\epsilon}_x^{(2)} \\
A_2^2 \boldsymbol{p}^2 + M\mathbf{w} = \boldsymbol{b}_2 & & \boldsymbol{\phi}_{xv} \\
\mathrm{diag}\big(\begin{pmatrix} A_2^2 & M \end{pmatrix}\begin{pmatrix} X^2 & Y^2 \\ Y^{2\top} & X^w \end{pmatrix}\begin{pmatrix} A_2^2 & M \end{pmatrix}^\top\big) = \boldsymbol{b}_2 \circ \boldsymbol{b}_2 & & \boldsymbol{\epsilon}_{xv} \\
\mathrm{diag}(A_3 \begin{pmatrix} X^1 & Y^{12} \\ Y^{12\top} & X^2 \end{pmatrix} A_4^\top) = \boldsymbol{0} & & \boldsymbol{\lambda} \\
p_j^1 = X_{jj}^1, & \forall j \in \mathcal{B}_1 & \boldsymbol{\psi}_{x1} \\
p_j^2 = X_{jj}^2, & \forall j \in \mathcal{B}_2 & \boldsymbol{\psi}_{x2} \\
 & & \\
\underline{\textit{Constraints on Random Parameters}} & & \\
M_1 \boldsymbol{w} = \boldsymbol{b} & & \boldsymbol{\phi}_v \\
\mathrm{diag}(M_1 X^w M_1^\top) = \boldsymbol{b} \circ \boldsymbol{b} & & \boldsymbol{\epsilon}_v \\
w_i = X_{ii}^w, & \forall i \in \mathcal{V} & \boldsymbol{\psi}_v \\
\boldsymbol{w} = \mu^v & & \boldsymbol{\nu} \\
X^w = \Sigma^v & & \Theta \\
 & & \\
CP = \begin{pmatrix} 1 & \boldsymbol{w}^\top & \boldsymbol{p}^{2\top} & \boldsymbol{p}^{1\top} \\ \boldsymbol{w} & X^w & Y^{2\top} & Y^{1\top} \\ \boldsymbol{p}^2 & Y^2 & X^2 & X^{12\top} \\ \boldsymbol{p}^1 & Y^1 & X^{12} & X^1 \end{pmatrix} \succcurlyeq_{cp} 0 & & \rho
\end{array}
$$

Set $\boldsymbol{\nu} = \boldsymbol{0}, \boldsymbol{\psi}_v = \boldsymbol{0}, \boldsymbol{\phi}_v = \boldsymbol{0}, \boldsymbol{\phi}_{xv} = \boldsymbol{0}, \boldsymbol{\psi}_{x1} = \boldsymbol{0}, \boldsymbol{\psi}_{x2} = \boldsymbol{0}, \boldsymbol{\phi}_x^{(1)} = \boldsymbol{0}, \boldsymbol{\phi}_x^{(2)} = \boldsymbol{0}, \Theta = \rho\mathbb{I}, \boldsymbol{\lambda} = \boldsymbol{0}, \boldsymbol{\epsilon}_{xv} = \rho\boldsymbol{1}, \boldsymbol{\epsilon}_v = \rho\boldsymbol{1}$, $\boldsymbol{\epsilon}_x^{(1)} = \boldsymbol{1}$ and $\boldsymbol{\epsilon}_x^{(2)} = \rho\boldsymbol{1}$ in the copositive program. Note that $\mathbb{I}$ denote the identity matrix. Then we have

$$CO_0 = \begin{pmatrix} \rho & \boldsymbol{0}^\top & \boldsymbol{0}^\top & \frac{1}{2}(-\hat{\boldsymbol{c}}_1)^\top \\ \boldsymbol{0} & \rho(\mathbb{I}+M^\top M + M_1^\top M_1) & \rho M^\top A_2^2 & -\frac{\hat{C}_2}{2} \\ \boldsymbol{0} & \rho(M^\top A_2^2)^\top & \rho(A_2^{2\top} A_2^2 + A_1^{2\top} A_1^2) & -\frac{\hat{C}_3}{2} \\ \frac{1}{2}(-\hat{\boldsymbol{c}}_1) & (-\frac{\hat{C}_2}{2})^\top & -\frac{\hat{C}_3^\top}{2} & A_1^{1\top} A_1^1 \end{pmatrix}$$

According to copostive Schur complement lemma, if $\begin{pmatrix} \mathbb{I}+M^{\top}M+M_1^{\top}M_1 & M^{\top}A_2^2 \\ (M^{\top}A_2^2)^{\top} & A_2^{2\top}A_2^2+A_1^{2\top}A_1^2 \end{pmatrix} \succ 0$, CO_0 is a strictly copositive matrix under the condition that $A_1^{1\top}A_1^1 \succ_{co} 0$. In the following, we show that the condition $A_2^{2\top}A_2^2+A_1^{2\top}A_1^2 \succ 0$ can guarantee this condition.

Consider an arbitrary vector $\begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix}$,

$$\begin{aligned} &\begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix}^{\top} \begin{pmatrix} \mathbb{I}+M^{\top}M+M_1^{\top}M_1 & M^{\top}A_2^2 \\ (M^{\top}A_2^2)^{\top} & A_2^{2\top}A_2^2+A_1^{2\top}A_1^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix} \\ =\ & \boldsymbol{v}_1^{\top}(\mathbb{I}+M_1^{\top}M_1)\boldsymbol{v}_1+\boldsymbol{v}_1^{\top}M^{\top}M\boldsymbol{v}_1+\boldsymbol{v}_1^{\top}M^{\top}A_2^2\boldsymbol{v}_2+\boldsymbol{v}_2^{\top}A_2^{2\top}M\boldsymbol{v}_1+\boldsymbol{v}_2^{\top}A_2^{2\top}A_2^2\boldsymbol{v}_2+\boldsymbol{v}_2^{\top}A_1^{2\top}A_1^2\boldsymbol{v}_2 \end{aligned} \tag{EC.1}$$

Notice that $\boldsymbol{v}_1^{\top}M^{\top}M\boldsymbol{v}_1+\boldsymbol{v}_1^{\top}M^{\top}A_2^2\boldsymbol{v}_2+\boldsymbol{v}_2^{\top}A_2^{2\top}M\boldsymbol{v}_1+\boldsymbol{v}_2^{\top}A_2^{2\top}A_2^2\boldsymbol{v}_2=(A_2^2\boldsymbol{v}_2+M\boldsymbol{v}_1)^{\top}(A_2^2\boldsymbol{v}_2+M\boldsymbol{v}_1)\geq 0$, $A_1^{2\top}A_1^2\succeq 0$ implies $\boldsymbol{v}_2^{\top}A_1^{2\top}A_1^2\boldsymbol{v}_2\geq 0$ and $\mathbb{I}+M_1^{\top}M_1\succ 0$ implies $\boldsymbol{v}_1^{\top}(\mathbb{I}+M_1^{\top}M_1)\boldsymbol{v}_1>0$ for any $\boldsymbol{v}_1\neq\mathbf{0}$. Hence for any vectors $\begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \end{pmatrix}$ with $\boldsymbol{v}_1\neq\mathbf{0}$, $(EC.1)>0$. On the other hand, if $\boldsymbol{v}_1=\mathbf{0}$,

$$(EC.1)=\boldsymbol{v}_2^{\top}A_2^{2\top}A_2^2\boldsymbol{v}_2+\boldsymbol{v}_2^{\top}A_1^{2\top}A_1^2\boldsymbol{v}_2=\boldsymbol{v}_2^{\top}(A_2^{2\top}A_2^2+A_1^{2\top}A_1^2)\boldsymbol{v}_2>0$$

since $A_2^{2\top}A_2^2+A_1^{2\top}A_1^2\succ 0$. Q.E.D.

Proof of Lemma 3. Denote the optimal dual variables to the dual problem (15) as $(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*,\boldsymbol{\delta}^*)$. Consider the following lost sale problem with zero inventory when there is no disruption. In order for the supply chain to sustain normal operation, the lost sale in this case should be 0, i.e. the optimal value in (EC.2) is 0.

$$\begin{aligned} Z(\boldsymbol{v},\boldsymbol{r})= \min_{(x_{ij},\boldsymbol{u},\boldsymbol{l})}\ & \sum_{j\in\mathcal{N}} f_j l_j \\ s.t.\ & \sum_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij}+l_j\geq d_j, && \forall j\in\mathcal{N} \\ & \sum_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij}-u_i\leq 0, && \forall i\in\mathcal{M} \\ & \sum_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij}I_{it}^{PT}}{B_{tj}}-u_j\geq 0, && \forall j\in\mathcal{M},t\in\mathcal{T}_j \\ & \sum_{i\in\mathcal{A}_k} u_i\leq c_k, && \forall k\in\mathcal{P} \\ & x_{ij}\geq 0,\boldsymbol{u},\boldsymbol{l}\geq\mathbf{0} \end{aligned} \tag{EC.2}$$

Its corresponding dual problem is

$$\begin{aligned} \max\ & \sum_{j\in\mathcal{N}} d_j\alpha_j-\sum_{k\in\mathcal{P}} c_k\delta_k \\ s.t.\ & (\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{s^1},\boldsymbol{s^2},\boldsymbol{s^3},\boldsymbol{s^4},\boldsymbol{s^5})\in\mathcal{F} \end{aligned} \tag{EC.3}$$

Let $(\boldsymbol{\alpha}^0,\boldsymbol{\beta}^0,\boldsymbol{\gamma}^0,\boldsymbol{\delta}^0)$ denote the optimal dual variable of Problem (EC.3). Due to LP strong duality, this dual problem also has optimal value of 0, i.e. $\sum_{j\in\mathcal{N}} d_j\alpha_j^0-\sum_{k\in\mathcal{P}} c_k\delta_k^0=0$. Furthermore, since the feasible region of Problem (15) coincide with that of Problem (EC.3), $(\boldsymbol{\alpha}^*,\boldsymbol{\beta}^*,\boldsymbol{\gamma}^*,\boldsymbol{\delta}^*)$ is also a feasible solution to Problem (EC.3). Therefore, we have $\sum_{j\in\mathcal{N}} d_j\alpha_j^*-\sum_{k\in\mathcal{P}} c_k\delta_k^*\leq\sum_{j\in\mathcal{N}} d_j\alpha_j^0-\sum_{k\in\mathcal{P}} c_k\delta_k^0=0$. Q.E.D.

Proof of Lemma 4.

Notice the structure of the feasible region of Problem (15) ($\mathcal{F}$ in EC. 2). For any feasible solution $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{s^1}, \boldsymbol{s^2}, \boldsymbol{s^3}, \boldsymbol{s^4}, \boldsymbol{s^5})$, we can always find $(\boldsymbol{\alpha}, K\boldsymbol{\beta}, K\boldsymbol{\gamma}, K\boldsymbol{\delta}, K\mathbf{s^1} + (K-1)\boldsymbol{\alpha}, K\mathbf{s^2}, K\mathbf{s^3}, K\mathbf{s^4}, \mathbf{s^5})$ feasible for arbitrary large K.

Notice in the objective of Problem (19), the coefficients of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are all non-positive. Hence, the optimal solution $\boldsymbol{\beta^*}, \boldsymbol{\delta^*}$ will not exceed the maximum value of $\boldsymbol{\alpha}$. We already know $\boldsymbol{\alpha} \leq \boldsymbol{f}$. Hence we can add valid cuts $\beta_j \leq \max_{i=1}^{n} f_i, \forall j = 1, \dots, m$, $\delta_k \leq \max_{i=1}^{n} f_i, \forall k = 1, \dots, p$. We also know that $\beta_j, \forall j$ and $\delta_k, \forall k$ has to be 0 if y is 0. We then have $\beta_j \leq \max_{i=1}^{n} f_i y, \forall j = 1, \dots, m$; $\delta_k \leq \max_{i=1}^{n} f_i y, \forall k = 1, \dots, p$.

Using a similar construction idea, we can see $(K\boldsymbol{\alpha}, K\boldsymbol{\beta}, K\boldsymbol{\gamma}, K\boldsymbol{\delta}, K\mathbf{s^1}, K\mathbf{s^2}, K\mathbf{s^3}, K\mathbf{s^4}, \mathbf{s^5})$ is always a feasible solution to Problem (19) as long as $K\boldsymbol{\alpha} \leq \mathbf{f}$. Therefore, one can verify that $\alpha_i \in \{0, f_i\}$ holds in optimal. In other words, $\alpha_i s_i^5 = 0, \forall i = 1, \dots, n$.

Next, we already have the constraint on $\boldsymbol{\alpha}$ that $\alpha_j + s_j^5 = f_j$. Since when $y = 0$, $\alpha_j = 0, \forall j$, it is valid to add the cut $\alpha_j + s_j^{12} = f_j y$.

Finally, $y \leq 1$ is valid due to the fact that y is binary. Q.E.D.

Proof of Lemma 5 We write down $A_1^1 \boldsymbol{x}_1 = \boldsymbol{b}_1^{(1)}$ specifically as follows

$$\left\{\begin{array}{ll}
\alpha_j - \beta_i + s_l^1 = 0, & \forall j \in \mathcal{N}, (i,j) \in \mathcal{G}_1 \\
-\beta_i + \sum\limits_{t \in \mathcal{T}_j} \frac{\gamma_t^j I_{it}^{PT}}{B_{tj}} + s_l^2 = 0, & \forall i \in \mathcal{M}, \forall j \in \mathcal{M}_2, (i,j) \in \mathcal{G} \\
-\delta_k + \beta_{i(k)} - \sum\limits_{t \in \mathcal{T}_{i(k)}} \gamma_t^{i(k)} + s_l^3 = 0, & \forall k \in \mathcal{P}, i(k) \in \mathcal{A}_k, i(k) \in \mathcal{M}_2 \\
-\delta_k + \beta_{i(k)} + s_l^4 = 0, & \forall k \in \mathcal{P}, i(k) \in \mathcal{A}_k, i(k) \in \mathcal{M}_1 \\
\alpha_j + s_j^5 = f_j & \forall j \in \mathcal{N} \\
\sum\limits_{j \in \mathcal{N}} d_j \alpha_j - \sum\limits_{k \in \mathcal{P}} c_k \delta_k + s^7 = 0 & \\
\beta_j + s_j^8 = \left(\max\limits_{i=1}^{n} f_i\right) y, \forall j = 1, \dots, m & \\
\delta_k + s_k^9 = \left(\max\limits_{i=1}^{n} f_i\right) y, \forall k = 1, \dots, p & \\
y + s^{11} = 1 & \\
\alpha_j + s_j^{12} = f_j y & \\
\boldsymbol{s^8}, \boldsymbol{s^9}, s^{11}, \boldsymbol{s^{12}} \geq \mathbf{0} & \\
\boldsymbol{\alpha} \in \mathbb{R}_+^n, \boldsymbol{\beta} \in \mathbb{R}_+^m, \boldsymbol{\gamma} \in \mathbb{R}_+^{tp}, \boldsymbol{\delta} \in \mathbb{R}_+^p & \\
\boldsymbol{s^1} \in \mathbb{R}_+^{|\mathcal{G}_2|}, \boldsymbol{s^2} \in \mathbb{R}_+^{|\mathcal{G}_1|}, \boldsymbol{s^3} \in \mathbb{R}_+^{m_2} & \\
\boldsymbol{s^4} \in \mathbb{R}_+^{m_1}, \boldsymbol{s^5} \in \mathbb{R}_+^n, s^7 \geq 0 &
\end{array}\right\} \tag{EC.4}$$

Denote the dual variable of each constraint defined by $A_1^1 \mathbf{x} = \mathbf{b_1}^{(1)}$ as $x_{ij}, \forall j \in \mathcal{N}, (i,j) \in \mathcal{G}$, $x_{ij}, \forall i \in \mathcal{M}, \forall j \in \mathcal{M}_2, (i,j) \in \mathcal{G}$, $\mathbf{u} \in \mathbb{R}^M$, $\mathbf{l} \in \mathbb{R}^N$ and $\boldsymbol{\theta}^{(i)}$ for $i = 1, \dots, 5$. We use $\mathbf{y}$ to represent all the dual variables defined above and then we can write down $A_1^{1\intercal}\mathbf{y}$ explicitly as

$$A_1^{1\intercal}\mathbf{y}=\left\{\begin{array}{lll}\sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij}+l_j+d_j\theta^{(1)}+\theta_j^{(5)}, & \forall j\in\mathcal{N} & (\alpha_j)\\ -\sum\limits_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij}+u_i+\theta_i^{(2)}, & \forall i\in\mathcal{M} & (\beta_i)\\ \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij}I_{it}^{PT}}{B_{tj}}-u_j, & \forall j\in\mathcal{M},t\in\mathcal{T}_j & (\gamma_j)\\ -\sum\limits_{i\in\mathcal{A}_k} u_i-c_k\theta^{(1)}+\theta_k^{(3)}, & \forall k\in\mathcal{P} & (\delta_k)\\ -f\sum\limits_{j\in\mathcal{M}}\theta_j^{(2)}-f\sum\limits_{k\in\mathcal{P}}\theta_k^{(3)}+\theta^{(4)}-\sum\limits_{j\in\mathcal{N}}f_j\theta_j^{(5)} & & (y)\\ \begin{pmatrix}\boldsymbol{x}^\intercal & \boldsymbol{u}^\intercal & \boldsymbol{l}^\intercal & \boldsymbol{\theta}^{(1)\intercal} & \dots & \boldsymbol{\theta}^{(5)\intercal}\end{pmatrix}^\intercal & & (\boldsymbol{s})\end{array}\right\}$$

where $f=\max\limits_{i=1}^{n} f_i$. Consider $\boldsymbol{l}=\epsilon\mathbf{1}$, $\boldsymbol{\theta}^{(5)}=\epsilon\mathbf{1}$, $\theta^{(1)}=\epsilon$, $\theta_i^{(2)}=\epsilon$, $\theta_k^{(3)}=c_k+2c_k\epsilon$, $\theta^{(4)}=Mf\epsilon+f(2\epsilon+1)\sum\limits_{k\in\mathcal{P}}c_k+\sum\limits_{j\in\mathcal{N}}f_j\epsilon+\epsilon$, where $\epsilon>0$. Notice that we assume there exists an interior point in the linear program (6) when there is no disruption and zero inventory in the supply chain. In other words, there exists strictly positive x_{ij}^0 and $\boldsymbol{u}^0$ such that

$$\left\{\begin{array}{ll}\sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} x_{ij}^0>d_j, & \forall j\in\mathcal{N}\\ -\sum\limits_{j\in\mathcal{M}\cup\mathcal{N},(i,j)\in\mathcal{G}} x_{ij}^0+u_i^0>0, & \forall i\in\mathcal{M}\\ \sum\limits_{i\in\mathcal{M},(i,j)\in\mathcal{G}} \frac{x_{ij}^0I_{it}^{PT}}{B_{tj}}-u_j^0>0, & \forall j\in\mathcal{M},t\in\mathcal{T}_j\\ \sum\limits_{i\in\mathcal{A}_k} u_i^0<(1+\epsilon)c_k, & \forall k\in\mathcal{P}\\ x_{ij}^0>0,\boldsymbol{u}^0>\mathbf{0} & \end{array}\right\}$$

Therefore, $A_1^{1\intercal}\begin{pmatrix}\boldsymbol{x}^0\\ \boldsymbol{u}^0\\ \boldsymbol{l}\\ \boldsymbol{\theta}\end{pmatrix}>0$, i.e. there exists a $\boldsymbol{y}_0$ such that $A_1^{1\intercal}\mathbf{y}_0>\mathbf{0}$. According to Lemma 3 in Hanasusanto and Kuhn (2017), we prove the claim. Q.E.D.

Proof of Lemma 6.

Recall that by Theorem 1, we have an equivalent completely positive reformulation (**??**) of $Z^m(\boldsymbol{r},\theta)=\max_{\tilde{v}\sim(\boldsymbol{\mu}^v,\Sigma^v)}\boldsymbol{E}[(Z(\tilde{\boldsymbol{v}},r)-\theta)^+]$. If we set $\boldsymbol{r}=\boldsymbol{r^*}$ and $\theta=0$, we have $Z^m(\boldsymbol{r^*},0)=\max_{\tilde{v}\sim(\boldsymbol{\mu}^v,\Sigma^v)}\boldsymbol{E}[(Z(\tilde{\boldsymbol{v}},r^*))^+]$. Since the lost sales are always greater than 0, we have $(Z(\tilde{\boldsymbol{v}},r^*))^+=Z(\tilde{\boldsymbol{v}},r^*)$. Therefore, we get $Z^m(\boldsymbol{r^*},0)=\max_{\tilde{v}\sim(\boldsymbol{\mu}^v,\Sigma^v)}\boldsymbol{E}[Z(\tilde{\boldsymbol{v}},r^*)]=Z^{\hat{e}}(\boldsymbol{r^*})$. Q.E.D.

Proof of Proposition 4 to Proposition 6

Suppose $(\boldsymbol{p}^{\boldsymbol{x}*},\boldsymbol{p}^{\boldsymbol{w}*},\boldsymbol{p}^{\boldsymbol{s}*},X^{x*},X^{w*},X^{s*},Y^{x*},Y^{xs*},Y^{ws*})$ is the optimal solution of Problem (23). Consider the decomposition of the optimal completely positive matrix.

$$\begin{pmatrix}1 & \boldsymbol{p}^{x*\intercal} & \boldsymbol{p}^{w*\intercal} & \boldsymbol{p}^{\boldsymbol{s}*\intercal}\\ \boldsymbol{p}^{x*} & X^{x*} & Y^{x*} & Y^{xs*}\\ \boldsymbol{p}^{\boldsymbol{w}*} & Y^{x*\intercal} & X^{w*} & Y^{ws*}\\ \boldsymbol{p}^{\boldsymbol{s}*} & Y^{xs*\intercal} & Y^{ws*\intercal} & X^{s*}\end{pmatrix}=\sum_{k\in\kappa_+}\zeta_k^{*2}\begin{pmatrix}1\\ \frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}\\ \frac{\hat{\boldsymbol{v}}_{\boldsymbol{k}}^*}{\zeta_k^*}\\ \frac{\hat{\boldsymbol{s}}_{\boldsymbol{k}}^*}{\zeta_k^*}\end{pmatrix}\begin{pmatrix}1\\ \frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}\\ \frac{\hat{\boldsymbol{v}}_{\boldsymbol{k}}^*}{\zeta_k^*}\\ \frac{\hat{\boldsymbol{s}}_{\boldsymbol{k}}^*}{\zeta_k^*}\end{pmatrix}^\intercal \tag{EC.5}$$

Where $\zeta_k^{*2}, \forall k \in \kappa_+$ specifies the worst-case probability of each scenario. The proof in the following is to construct feasible solutions to the completely positive program (23) when a certain parameter changes based on $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}, \forall k \in \kappa_+$.

Proof of Proposition 4.

With a bit abuse of notation, for each scenario k, notice that $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$ indeed corresponds to the dual and slack variables $\{\boldsymbol{\alpha}_k, \boldsymbol{\beta}_k, \boldsymbol{\gamma}_k, \boldsymbol{\delta}_k, T_k^R, y_k, \boldsymbol{s^j}_k (\forall j \neq 6), s_k^6\}$. If T_t^r decreases by ϵ, i.e., $T_t^r \leftarrow T_t^r - \epsilon$. We consider following three cases:

- For those k such that $\frac{\hat{v}_{kt}^*}{\zeta_k^*} = 1$, $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$ will remain feasible after changing T_t^r to $T_t^r - \epsilon$.
- For those k such that $\frac{\hat{v}_{kt}^*}{\zeta_k^*} = 0$ and $T_t^r < T^R(\boldsymbol{v})$, then construct $\hat{s}_{tk}^6 := s_{tk}^{6*} + \epsilon$ and the other variables remain the same values as $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$. Then the new solution, denoted as $\frac{\hat{\boldsymbol{\tau}}_k}{\zeta_k^*}$, is feasible under the new parameter $T_t^r - \epsilon$, with the objective value decreased by $\epsilon c_t \delta_t^*$.
- For those k such that $\frac{\hat{v}_{kt}^*}{\zeta_k^*} = 0$ and $T_t^r = T^R(\boldsymbol{v})$, if vendor t attains the maximum TTR alone (ϵ is small enough such that t's TTR is still the largest among all), then we construct a new solution with $\hat{T}_k^R := T^R - \epsilon$, $\hat{s}_{ik}^6 := s_{ik}^{6*} - \epsilon \geq 0, \forall i \neq t$ and $\hat{s}_{tk}^6 := s_{tk}^{6*}$. The objective value is decreased by $\epsilon(\sum_{j \in \mathcal{N}} \alpha_j(\boldsymbol{v}) d_j - \sum_{i \neq t} c_i \delta_i(\boldsymbol{v}))$, which is no less than $\epsilon c_t \delta_t(\boldsymbol{v})$ according to Lemma 3.
- For those k such that $\frac{\hat{v}_{kt}^*}{\zeta_k^*} = 0$ and $T_t^r = T^R(\boldsymbol{v})$ but there are other vendors attaining the maximum TTR together with vendor t. Then let $\hat{s}_{tk}^6 := s_{tk}^{6*} + \epsilon$ and the other variables remain the same values as $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$. It would be a feasible solution to the problem when T_t^r decreases by ϵ. In this case, the objective value is decreased by $\epsilon c_t \delta_t^*$.

In summary, we have $Z^{\hat{e}}(\boldsymbol{r}^*, \frac{\hat{\boldsymbol{v}}_{\boldsymbol{k}}^*}{\zeta_k^*})|_{T^r-\epsilon} - Z^{\hat{e}}(\boldsymbol{r}^*, \frac{\hat{\boldsymbol{v}}_{\boldsymbol{k}}^*}{\zeta_k^*})|_{T^r} \geq -\epsilon c_t \delta_t^* (1 - \frac{\hat{v}_{kt}^*}{\zeta_k^*}), \forall k \in \kappa_+$.

Then the corresponding completely positive matrix

$$\begin{pmatrix} 1 & \boldsymbol{p}^{x'\top} & \boldsymbol{p}^{w'\top} & \boldsymbol{p}^{\boldsymbol{s}'\top} \\ \boldsymbol{p}^{x'} & X^{x'} & Y^{x'} & Y^{xs'} \\ \boldsymbol{p}^{\boldsymbol{w}'} & Y^{x'\top} & X^{w'} & Y^{ws'} \\ \boldsymbol{p}^{\boldsymbol{s}'} & Y^{xs'\top} & Y^{ws'\top} & X^{s'} \end{pmatrix} = \sum_{k \in \kappa_+} \begin{pmatrix} \zeta_k^* \\ \hat{\boldsymbol{\tau}}_k \\ \hat{\boldsymbol{v}}_{\boldsymbol{k}}^* \\ \hat{\boldsymbol{s}}_{\boldsymbol{k}}^* \end{pmatrix} \begin{pmatrix} \zeta_k^* \\ \hat{\boldsymbol{\tau}}_k \\ \hat{\boldsymbol{v}}_{\boldsymbol{k}}^* \\ \hat{\boldsymbol{s}}_{\boldsymbol{k}}^* \end{pmatrix}^\top \tag{EC.6}$$

is a feasible solution to (23) with $Z^{\hat{e}}(\boldsymbol{r}^*)|_{T^r-\epsilon} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{T^r} \geq -\epsilon c_t (p_{\delta t}^* - \epsilon Y_{\delta tt}^*)$. Therefore, $\lim_{\epsilon \to 0^+} \frac{Z^{\hat{e}}(\boldsymbol{r}^*)|_{T^r} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{T^r-\epsilon}}{\epsilon} \leq c_t (p_{\delta t}^* - Y_{\delta tt}^*)$. Q.E.D.

Proof of Proposition 5.

This proof follows the exactly same logic as the one for Proposition 4, we omit the details here.

Proof of Proposition 6.

It is easy to see the completely positive program (23) is convex in $r_i, i = 1, ..., m$. Consider decrease r_t by $\epsilon \geq 0$, i.e., $r_t \leftarrow r_t - \epsilon$, $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$ will remain feasible. Therefore $Z^{\hat{e}}(\hat{\boldsymbol{r}^*})|_{r_t-\epsilon} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{r_t} \geq \epsilon \boldsymbol{p}_{\beta t}^*$. In contrast, consider increasing r_t by $\epsilon \geq 0$, i.e., $r_t \leftarrow r_t + \epsilon$, $\frac{\boldsymbol{\tau}_k^*}{\zeta_k^*}$ will remain feasible. Therefore $Z^{\hat{e}}(\hat{\boldsymbol{r}^*})|_{r_t+\epsilon} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{r_t} \geq -\epsilon \boldsymbol{p}_{\beta t}^*$. In summary, we have $\lim_{\epsilon \to 0^+} \frac{Z^{\hat{e}}(\hat{\boldsymbol{r}^*})|_{r_t+\epsilon} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{r_t}}{\epsilon} \geq -\boldsymbol{p}_{\beta t}^*$, $\lim_{\epsilon \to 0^+} \frac{Z^{\hat{e}}(\hat{\boldsymbol{r}^*})|_{r_t} - Z^{\hat{e}}(\boldsymbol{r}^*)|_{r_t-\epsilon}}{\epsilon} \leq -\boldsymbol{p}_{\beta t}^*$. Hence $\frac{\partial Z^{\hat{e}}(\boldsymbol{r}^*)}{\partial r_t} = -p_{\boldsymbol{\beta} t}^*$. Q.E.D.