Modeling Pharmaceutical Supply Chains to Mitigate Drug Shortages

by

Emily L. Tucker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Industrial and Operations Engineering)
in the University of Michigan
2020

Doctoral Committee:

Professor Mark S. Daskin, Chair
Professor Xiuli Chao
Professor Wallace J. Hopp
Associate Professor Jun Li
Professor Burgunda V. Sweet

Emily L. Tucker

eltuck@umich.edu

ORCID iD: 0000-0003-1370-1642

ACKNOWLEDGEMENTS

I am profoundly grateful for the many people who have supported me throughout my PhD. First, to my family – you are the greatest. A very special thank you to my parents Susan and Mike and my brother Michael.

To my advisor – Mark Daskin, thank you. I appreciate your support of my PhD work and of my career. You have opened many doors. To my dissertation committee – Wally Hopp, Gundy Sweet, Xiuli Chao, and Jun Li. Thank you for the many productive conversations and collaborations. My work has been strengthened by many other collaborators and experts as well – notably Kayse Maass, Erin Fox, Allan Pfitzenmaier, Matthew Rosenberg, Abby Smith, Peg Hill-Callahan, Steve Gill, Bob Merion, and Alan Leichtman.

Thank you to the Daskin research group – Kayse Maass, Ece Sancı, Nima Salehi, and Lauren Czerniak. I appreciate your mentorship, support, and friendship. Thank you to the many undergraduates I have had the opportunity to work with: Abbey Weis, Sherry Li, Yizhou Cao, Madeleine Videira, Alex Lopz, Deanna Handley, Hannah Strat, Sabrina Cottrell, Evelyn Gendron, Hannah Schapiro, Jordan Schevil, Erica Segre, and Eva Cahnman.

To the entire IOE department – thank you to the many faculty, staff, and students who taught classes, offered advice, and supported me in this journey. I am convinced the PhD cohort that started in the Fall of 2015 is the best ever. Special thanks to Karmel Shehadeh, Ece Sancı, and Niusha Navidi for being great friends. Thank you to the IOE staff, especially Rebekah Smith, Cathy Boblitt, Valerie Martin, Tina Picano Sroka, Gwen Brown, and Amanda Godwin for making this process far smoother.

My PhD was funded by many sources – the National Science Foundation through a Graduate Research Fellowship, a Rackham Merit Fellowship, and the MCubed program. Thank you for the opportunity to do this work.

Thank you to the many people in Ann Arbor who have made this place home. I have made some of the best friends I could ever ask for – Pétra Vande Zande, Emma Davis, Emily Johengen, Anna Michmerhuizen, Dayna Asante-Appiah, and Dianne Roeper. Thank you for all you've done for me. GradCru has been spring of joy and friendship – thank you to everyone who has been involved. Tuesday dinners have been a highlight of my weeks – thank you to George and Mary Lindquist and the many other helping hands. Thank you to my family at Grace Bible Church. It is a pleasure to worship and serve with y'all.

To my friends who have stuck with me for so many years – including many from middle and high school (especially Jing Yuan, Lin Yuan, Sarah Paleg, and Julia Duke) and college (Gretchen Stokes and Sarah Shelton among so many others!) who have inspired and supported me for years – thank you. In just one example, when I started my PhD, Gretch earnestly promised that she would read my entire dissertation – a great show of support that she now probably regrets upon learning its final length.

To so many people who helped me get to where I am today – through conversations, connections, and love – thank you. This group includes the NC State folks who introduced me to IE and encouraged me to pursue a PhD (Julie Ivy, Thom Hodgson, and Anita Vila-Parrish); the Park Scholarship program which invested so much in me; and my wonderful colleagues and friends at RTI who taught me how to be a researcher (especially Katherine Hicks).

And most of all to Jesus Christ. He is my ultimate hope, source of all joy, and holds all things together (Colossians 1:17).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The use of medically-necessary drugs has extended the lives of countless patients. While healthcare providers rely on the pharmaceutical industry for treatments, in recent years, the drug supply in the United States has become volatile and drug shortages are common. Shortages are considered a public health crisis and are often caused by disruptions to vulnerable pharmaceutical supply chains. The tightly optimized supply chains have little redundancy and low levels of inventory. This combination can cause minor supply interruptions to become widespread shortages. I study the dynamics of shortages by developing new models of supply chains under disruption, and I identify regulations and incentives to induce companies to reduce the occurrence and impact of shortages. There has been minimal analysis on the quantitative impact of proposed policies.

I present four mathematical models. The first two are static supply chain design problems (SCDD and SCDD-I). The company decides at the beginning of the horizon how to configure its supply chain. In the second model (SCDD-I), the company may also choose to hold inventory. The models are two of the first to include disruptions and recovery over time. They are solved using Sample Average Approximation. The analyses suggest that it is either not economically feasible or attractive for companies to maintain resilient supply chains for some drugs that are vulnerable to shortage. I use the models to compare policies that have been proposed to reduce shortages. It is less expensive to raise prices in combination with resilience requirements than to raise prices alone. Requiring a second supplier may have the largest incremental benefit than requiring a back-up at other levels of the supply chain.

The third model (D-SCDD) is a dynamic supply chain design model. It is a multi-stage stochastic program. At the beginning of the time horizon, the company selects the supply chain configuration and may add components or stop production if disruptions occur. The formulation applies the geometrically-distributed times to recover and disruption via an inverse sampling approach to maintain stage-wise independence. It is solved using the Stochastic Dual Dynamic Integer Programming (SDDiP) algorithm. I find that substantial reductions in the lead times to add components or reducing the mean time to recover disrupted components may reduce shortages. Minor lead time reductions have little impact.

The fourth model (SCR) is comprised of closed-form expressions that describe the reliability characteristics of a given supply chain configuration. The analyses provide evidence that increasing component quality would be effective at reducing shortages. The model can also be used to calculate break-even prices.

This project provides insight to policymakers and companies to support the profitable production of a reliable drug supply. The implications and use of these models is widely relevant to other regulated industries and supply chains under distribution.

CHAPTER I

Introduction

--

"How do you tell a patient who has recently received a diagnosis of cancer that the first-choice chemotherapy agent is not currently available for his treatment?" (Berry 2014)

"All ask the same question: 'How could this happen in the United States of America?'" (Kweder and Dill 2013)

--

Drug shortages regularly affect health care in the United States (US). They are often caused by disruptions to fragile pharmaceutical supply chains (GAO 2016). The factors that contribute to these disruptions are complex, and shortages continue to persist. There has been very limited work to analyze policy proposals quantitatively. Existing mathematical models that consider disruptions are not able to handle the nuances of shortages in the pharmaceutical industry.

In this dissertation, I address two goals:

- To develop new mathematical models that capture the relevant characteristics of disruption-related decisions in the pharmaceutical industry

- To provide evidence to support policy discussions with the aim of reducing drug shortages

In the remainder of this introduction, I present an overview of drug shortages in the US and an outline of the technical chapters.

1. Background on drug shortages

Shortages became widespread in the US beginning in 2010, and in a 2011 survey, 99.5% of hospitals reported at least one drug shortage in the six months prior to the survey (American Hospital Association 2011). In 2018, all 719 respondents (hospital pharmacy managers and leaders) reported they had experienced at least one drug shortage in the prior year, and 69.2% reported more than 50 (Hantel et al., 2019). Among oncology drugs, 34% of marketed drugs were short at least once between 2001-2017 (Hosseini et al., 2018). Estimates projected that 550,000 patients were affected by an oncology drug shortage in 2011 (IMS 2011). Saline shortages were projected to have affected millions of patients since 2003 (Chen et al., 2016). Currently, the COVID-19 (coronavirus) outbreak has begun to affect drug supply chains with factory closures in China (McCarthy 2020). The long term effects on drug manufacturing are unclear though concerning.

Shortages tend to affect drugs that are generic and/or sterile injectable (Dill and Ahn 2014, Le et al. 2011). These types of drugs tend to have higher production costs and lower prices. Drugs vulnerable to shortage are not necessarily representative of drugs as a whole. For many drugs, high prices are a major concern; those are not necessarily the drugs affected by shortages (Frakt 2016).

Most drug shortages – when the source is reported – are caused by supply chain disruptions (GAO 2016). If interruptions occur, it is difficult to ramp up production. The

industry has moved to just-in-time inventory practices, particularly for generic drugs, thereby making drug supply chains more brittle. Low profit margins and prices may lead to inadequate investments in resiliency strategies and quality systems (Woodcock and Wosinska 2013).

Shortages have translated into poor health outcomes, and several patients have died. During the norepinephrine shortage, there was a 3.7% increase in inpatient mortality (35.9% to 39.6%) for patients with septic shock (Vail et al., 2017). For pediatric cancer patients, two-year event-free survival dropped during the mechlorethamine shortage (Metzger et al. 2012). Shortages are costly. The national cost in labor associated with mediating drug shortages was estimated in 2019 to be $359 million annually (Vizient 2019). Purchasing alternative medications, if available, is also expensive. The most recent estimate was $200 million in 2011 (Cherici et al., 2011).

2. Modeling

Drug shortages continue in high numbers, and mathematical modeling to tackle the crisis has been limited (Tucker et al., 2020a). The strategic decisions that pharmaceutical companies make about their supply chains have direct effects on shortages when disruptions occur and their impacts. There are open questions about which policies would be effective at reducing shortages; the effects of increasing supply chain adaptability; and evaluating the risk of shortages. To consider these questions, new supply chain models that include both disruption and recovery over time are needed. Throughout, I will refer to supply chain "components." These represent the key facilities within the supply chain where disruptions can lead to shortages, i.e., Active Pharmaceutical Ingredient (API) suppliers; manufacturing plants; and manufacturing lines (GAO 2016). A summary of the technical chapters is as follows, and an overview of the models and associated features is presented in Table 1.

Chapter 2 addresses the question of resilient supply chain design when disruptions and recovery can occur over time. The chapter presents two models of static supply chain design. In both models, the configuration decisions made in the initial stage are maintained throughout the remainder of the horizon. The first, the Supply Chain Design under Disruption (SCDD) model, is a two-stage stochastic program where the decision-maker selects the supply chain configuration in the initial stage, and facility availability is revealed in the second stage for each of the remaining time periods. Operational decisions are made, and demand may be met or there is a shortage. The second model is the Supply Chain Design under Disruption with Inventory (SCDD-I) model. This is a multi-stage stochastic program where the company selects the supply chain configuration and target safety stock level in the initial stage. In the following stages, facility availability is realized. The operational decisions – ordering raw materials, producing the drug, inventory replenishment, and amount of demand to meet – are made. An exogenous replenishment rule implies standard non-anticipativity constraints are redundant and simplifies the model structure. Both models are solved using Sample Average Approximation (SAA) (Kleywegt et al. 2002). The models are used to conduct extensive analysis on policies that could induce generic injectable oncology drug companies to be more resilient. Two example generic injectable oncology drugs are used as case examples (vinblastine sulfate and vincristine sulfate). The key contributions of this paper are as follows:

- It presents new supply chain design models that consider disruptions and recovery. They include the new combination of features: the inclusion of time; disruptions at multiple echelons and multiple components simultaneously; multiple resiliency strategies (redundancy and inventory).

- The model with inventory (SCDD-I) incorporates a replenishment rule that forces the standard non-anticipativity constraints to hold implicitly.
- The analyses shift the standard perspective of supply chain resiliency models. Rather than a company optimizing resiliency, they consider how to *induce* a profit-maximizing company to be resilient.
- It is the first paper to compare several policies that have been proposed to reduce drug shortages (including mandating redundancy or inventory; adding failure-to-supply clauses; and price increases).

This chapter has been published as Tucker et al. (2020b) and is co-authored with Mark Daskin, Burgunda Sweet, and Wallace Hopp.

Chapter 3 considers the question of a *dynamic* supply chain design. When disruptions occur, pharmaceutical companies have a limited ability to adjust their supply chain configurations. The process is limited by lengthy review times by the Food and Drug Administration (GAO 2016). Typically, disruption-related models assume that companies will remain in the market if a disruption occurs (Snyder et al., 2006). Their focus is rather on how to prevent disruptions or how to recover. This chapter considers different questions, namely i) how the lead times to add new components affect shortages and ii) how disruptions may affect a company with low profit-margin decisions to remain in the market. The analyses are run using an example generic injectable oncology drug (vincristine sulfate).

To address these questions, I introduce a new dynamic supply chain design model that incorporates disruptions and recovery. In the initial stage, the company selects its initial supply chain configuration that it maintains throughout the time horizon (unless it leaves the market). In the subsequent stages, facility availability is progressively revealed over time, and the company

may choose to maintain the initial configuration, to begin the process of adding new components, or to discontinue production and leave the market. In each stage, demand may be met or there is a shortage.

The model is solved using the Stochastic Dual Dynamic Integer Programming (SDDiP) algorithm (Zou et al. 2019). This algorithm requires the uncertainty to be stage-wise independent; yet disruptions and recovery within the pharmaceutical industry is geometrically distributed and follow different distributions (Chapter 2). To consider this, I reformulate the model to sample probability values from the inverse of the cumulative distribution function (CDF) of the appropriate distributions. Then the model applies the appropriate stochastic availability parameter via the constraints. This induces the realizations of uncertainty to be stage-wise independent and allows the SDDiP algorithm to be applied.

The key contributions are as follows:

- It is the first dynamic supply chain design model that considers disruptions and recovery.
- It applies multiple distributions for the times to disruption and recovery in a stage-wise independent approach.
- The analysis addresses a key public health question of whether reduced lead times to add components would reduce shortages.
- The model is used to study the relationship between disruptions and product discontinuations.

In Chapter 4, I present a new model of supply chain reliability. The previous chapters consider complex optimizations where the decision-maker seeks to design the pharmaceutical

supply chain. In this chapter, I consider the question of *evaluation* – what is the reliability of a given supply chain structure?

By considering an exogenous (given) supply chain structure, I develop closed-form equations of the probability of shortage, the average time between shortages, and the length of shortages when they occur. As the model is presented in closed-form, they could be solved using calculations within a spreadsheet or even by hand (in contrast to the specialized optimization software needed in Chapters 2 and 3). Pharmaceutical companies could use the model as they conduct internal risk analysis or by external regulators if they have access to key supply chain characteristics.

The model is used to evaluate the supply chain reliability of a generic injectable oncology drug (vincristine sulfate) under several conditions. Multiple potential configurations are considered. The analyses study the effects of reducing the expected time to recovery or reducing the expected disruption rate. I also consider profitability and costs. I calculate the prices at which more reliable supply chains become more profitable, and the prices where the company breaks even for different configurations.

The contributions of this chapter are:

- I present a new model for evaluating the reliability of pharmaceutical supply chains to disruptions. The equations are presented in closed-form.
- I analyze the effects of maintaining higher-quality components and quicker recovery after disruptions
- I conduct a pricing analysis to determine break-even pricing.

Table 1. Summary of models

		Features						
Model name	**Acronym**	**Closed-form solutions**	**Optimize supply chain**	**Includes inventory**	**Can change configuration**	**Model type**	**Solution method**	**Ch.**
Supply Chain Design under Disruption	SCDD		✓			Two-stage SP	SAA	2
Supply Chain Design under Disruption with Inventory	SCDD-I		✓	✓		Multi-stage SP	SAA	2
Dynamic Supply Chain Design under Disruption	D-SCDD		✓		✓	Multi-stage SP	SDDiP	3
Supply chain reliability	SCR	✓				Analytical	Closed-form; Simulation	4

Ch. = chapter; SAA = Sample Average Approximation; SDDiP = Stochastic Dual Dynamic Integer Programming; SP = stochastic program

The three technical chapters are framed as standalone papers. They each include introductions to the particular problem addressed, reviews of the relevant literature, methodology, results, and discussion. The conclusion in chapter 5 summarizes the dissertation and discusses future research directions.

CHAPTER II

Incentivizing Resilient Supply Chain Design to Prevent Drug Shortages: Policy Analysis Using Two- and Multi-Stage Stochastic Programs

1. Introduction

Over the past decade, the United States (US) has experienced unprecedented shortages of medically-necessary drugs. In 2015 alone, 427 drugs were unavailable (GAO 2016). Shortages last 14 months on average and can have negative effects on patient safety, clinical outcomes, and health system costs (GAO 2016, Tucker et al., 2020a). They are often caused by disruptions to non-resilient pharmaceutical supply chains. The question of how best to reduce their impact is pressing for patients and the US healthcare system.

Drugs that have been short span a variety of therapeutic classes including central nervous system (anesthetics), cardiovascular, anti-infective, and oncology agents (UUDIS 2016). When shortages occur, treatment of patients may be delayed, changed, or cancelled entirely (Goldsack et al. 2014, McLaughlin et al. 2013). Shortages have been associated with patient deaths (Fox et al. 2014, Vail et al. 2017), and managing them has been compared to dealing with a "natural disaster or national emergency… occur[ring] on a daily basis" (Fox et al. 2014).

Costs associated with managing shortages are high. It has been estimated that traditional US health systems spend $216-359 million each year on labor costs to manage shortages (Kaakeh et al. 2011, Vizient 2019) and $200 million annually to purchase substitute drugs and hold extra inventory (Fox et al. 2014).

Shortages are caused by a variety of factors. Production may be delayed, companies may leave the market, or production lines may be contaminated. In some cases, Food and Drug Administration (FDA) inspections have uncovered quality concerns at manufacturing and raw material facilities that have led to extended shutdowns (Fox et al. 2014, Palmer 2016). Some drugs have been short multiple times because of intermittent manufacturing issues (UUDIS 2016). Often companies do not report the direct cause of a shortage, but of those reported, 82% are caused by a supply chain disruption (GAO 2016).

There has been substantial research on mitigating disruptions in non-pharmaceutical supply chains, and researchers have found that maintaining some degree of resiliency is often optimal (Snyder et al. 2016, Tomlin 2006). However, shortages of drugs persist. This may be due, at least in part, to the unique challenges of the highly regulated pharmaceutical industry. Whenever a pharmaceutical company changes its supply chain, it must go through a lengthy FDA approval process (GAO 2016). This requirement makes it difficult to adapt to disruptions. When this is combined with the complex manufacturing processes of sterile drugs, there are inherent vulnerabilities. A stark contrast can be seen with other industries. When a fire shut down the sole supplier of a critical part for Toyota years ago, they were able to ramp-up other suppliers and resume production within days (Nishiguchi and Beaudet 1998). When Hurricane Maria shut down pharmaceutical manufacturing plants in Puerto Rico, drug shortages continued to affect the healthcare system for months (Gottlieb 2018b, Thomas and Kaplan 2017). The strategic resiliency decisions made prior to disruption become critical.

Currently, companies that manufacture drugs that are particularly vulnerable to shortages (e.g., sole-source generic injectables) hold little backup capacity and maintain little to no safety stock (Fox et al. 2014, GAO 2016, Woodcock and Wosinska 2013). Profit margins are low

(GAO 2016), and there are few consequences if there is a shortage. Companies rarely pay penalties if they cannot supply a drug, and the risk of losing market share to a new competitor is small because of high barriers to entry (GAO 2016, Haninger et al. 2011, Jia and Zhao 2017).

Yet, drugs affected by shortages are often medically-necessary and life-sustaining. The fundamental question becomes: how could pharmaceutical supply chains be strengthened to provide a reliable drug supply? I will answer this by considering two specific questions: i) is it optimal for companies to choose low resiliency? and ii) if so, what is the best way to induce companies to maintain resilient supply chains?

Several strategies have been proposed. These include regulatory changes, e.g., require companies to maintain redundancy (ASHP 2013, Chabner 2011, FDA 2013, Gehrett 2012, Health Policy Brief: Drug Shortages 2014, Jarosławski et al. 2017) or hold safety stock (ASHP 2013, Gupta and Huang 2013, Jarosławski et al. 2017, Wiggins et al. 2014). Others include contractual changes such as strengthening failure-to-supply clauses (Conti 2011, FDA 2013, Haninger et al. 2011, Health Policy Brief: Drug Shortages 2014, Jia and Zhao 2017, Reed et al. 2016) and increasing prices (Chabner 2011, Gatesman and Smith 2011, Health Policy Brief: Drug Shortages 2014, Link et al. 2012). Limited analyses of the potential effects of these proposals have been conducted despite calls from experts for such studies (FDA 2013, Fox et al. 2014, Fox and Tyler 2013, ISPE and Pew Charitable Trusts 2017, Roberts et al. 2012). It remains unclear whether market-based interventions or regulatory changes would be more effective.

In this chapter, I seek to fill this gap to consider why pharmaceutical companies may set up either vulnerable or resilient supply chains and to analyze how the proposed policy changes would affect supply chain decisions. As different policies have different implementation costs, I

analyze the social-efficiency – how to reduce shortages to a specified level for the lowest cost. The analyses are focused on generic, oncology drugs and include two steps: first, I develop new pharmaceutical supply chain design models, and second, I change the underlying market conditions to analyze the effects of the proposed policies.

The remainder of this chapter is organized as follows. In Section 2, I review the relevant literature and discuss contributions of the analysis. In Section 3, I present the base model that includes redundancy as a resiliency strategy, and in Section 4, I develop the extension that includes safety stock. I discuss the solution methods in Section 5. In Section 6, I present case examples of two oncology drug supply chains. In Section 7, I discuss the results and policy implications and conclude.

2. Literature review

This work relates to several streams of literature, and I briefly review relevant studies that focus on pharmaceuticals, supply chain risk management, disruptions, and incentives.

2.1. Pharmaceutical modeling

The operations research and management science community has only recently begun to study drug shortages. Kim and Scott Morton (2015) analyzed factors that contribute to shortages with a game theory model of two competing manufacturers of perfectly substitutable generic injectable drugs. They suggested that spare capacities may have been removed when prices dropped in the early 2000s, revealing underlying vulnerabilities that led to shortages. In one of the only papers to evaluate policy, Jia and Zhao (2017) developed a model of contracts between key stakeholders to analyze the effects of failure-to-supply clauses and price increases. At the beginning of the contracting period, the manufacturer allocates production capacity at a single echelon and decides its inventory policy under stochastic supply and demand. The authors used

this framework to study case examples of fluorouracil, cytarabine, and bleomycin and found Pareto-improving contracts for each stakeholder. Others have studied inventory control for hospitals struggling with shortages (Saedi et al. 2016), inventory policies as a response to product recalls (Azghandi et al. 2018), and inventory policies related to human behavior and shortages (Doroudi et al. 2018). Jacobson, Sewell, and Proano (2006) analyzed the size of the Strategic National Stockpile of pediatric vaccines.

Optimization is regularly applied more broadly in the pharmaceutical literature (Narayana et al. 2014, Shah 2004) though work on supply chain design is uncommon. Exceptions include a multi-stage stochastic program which considered demand uncertainty (Guillén et al. 2006) and a four-echelon supply chain model under uncertainty in demand, cost, and desired safety stock levels (Mousazadeh et al. 2015). In this chapter, I consider supply chain design models where disruptions are a source of supply uncertainty.

2.2. Supply chain risk management

Supply chain risk management (SCRM) is a large research area that considers how companies structure and operate their supply chains to provide products to customers in the presence of uncertainty. There have been several reviews of this literature (including Ho et al. 2015, Tang 2006, Tang and Musa 2011), and one of the early reviews identified four main domains of SCRM – managing supply, demand, products, and information (Tang, 2006). In a seminal paper, Chopra and Sodhi (2004) discussed several strategies to manage risk, including capacity, inventory, redundancy, and flexibility.

Since then, many authors have developed quantitative models to analyze strategies in different contexts. Objectives have varied from cost minimization to bi-objective frameworks that trade off profit and risk (Nagurney 2006, Tomlin 2006). Some researchers explicitly

penalize unmet demand (Dada et al. 2007, Schmitt et al. 2010), and others consider robust approaches (O'Hanley and Church 2011). Researchers have found that supply and demand risks should be managed differently (Schmitt et al. 2015, Snyder and Shen 2006). When the risk is supply chain disruptions, Tomlin (2006) found that it is rarely optimal to passively accept risk; companies should nearly always select some level of resiliency. Yet pharmaceutical companies often passively accept risk for generic injectable drugs (Fox et al. 2014, GAO 2016, Woodcock and Wosinska 2013). One open question this chapter seeks to address is whether this choice is optimal.

2.3. Disruptions

Within the field of SCRM, many models have focused on disruptions as a source of supply-side risk. Snyder et al. (2016) presented an extensive review of this area. Common mitigation strategies include maintaining redundancy, holding inventory, or sourcing from multiple suppliers.

Redundancy or backup capacity decisions are often considered within the facility location literature. A number of studies have considered where to locate facilities at a single echelon given disruptions (e.g., Snyder and Daskin 2005). Fewer papers have considered decisions for multiple echelons (e.g., the robust approach of Peng et al. 2011).

Within the inventory literature, researchers have extended standard, single-supplier models to include disruptions. A survey of key models is available from Atan and Snyder (2012). In the Economic Order Quantity model with Disruptions (EOQD), the supplier is disrupted according to a continuous-time Markov chain (presented by Parlar and Berkin 1991; corrected by Berk and Arreola-Risa 1994). One extension includes the risk of disruptions at the retailer (Qi et al. 2009). In the periodic-review framework, Song and Zipkin (1996) proved that

a base-stock policy is optimal for a single echelon if the order costs are linear, and Schmitt, Snyder, and Shen (2010) derived the exact and approximate expected costs when there is stochastic demand.

Where there are multiple suppliers, inventory decisions can be made using extensions to the EOQD model (Gurler and Parlar 1997, Parlar and Perry 1996) or a network of queues (Song and Zipkin 2009). Schmitt and Tomlin (2012) analyzed whether single- or multi-sourcing is optimal in different contexts, and Saghafian and Van Oyen (2012) studied the effects of the flexibility of the backup. Mak and Shen (2012) considered dynamic sourcing to mitigate both demand and supply uncertainty.

Companies may also consider multiple disruption mitigation strategies. Tomlin (2006) compared a firm's decision to dual-source with holding inventory, rerouting, and passive acceptance. Others analyzed strategies for supply chain networks (Bundschuh et al. 2003, Hopp and Yin 2006, Schmitt 2011). These included separate models for multi-sourcing, safety stock, and meeting an expected service level (Bundschuh, Klabjan, and Thurston, 2003) and models that traded off backup capacity and safety stock (Hopp and Yin 2006, Schmitt 2011). Schmitt (2011) noted that inventory is helpful for shorter, more frequent disruptions, and backup capacity is better for less frequent, longer disruptions. MacKenzie, Barker, and Santos (2014) studied the decisions suppliers and firms make during and after a disruption, including whether to switch to an alternate facility. Where trade-offs were evaluated, the key metric was company cost or profit (e.g., Hopp and Yin 2006); this chapter focuses on the overall societal cost, i.e., social-efficiency.

Disruption models that included inventory often considered decisions over multiple time periods (e.g., Berk and Arreola-Risa 1994, Tomlin 2006), though location and design models generally did not. The latter often assumed the system returned to steady-state before another

disruption occurred (e.g., Hopp and Yin 2006; Kim and Scott Morton 2015; Schmitt 2011) or implicitly considered a single time period (e.g., Bundschuh et al. 2003). The strategic design models that included time considered decisions for a single echelon (Fattahi et al. 2017, Losada et al. 2012) or single layer of arcs between two echelons (Mak and Shen 2012). A recent review noted that in general very few supply chain design problems under uncertainty have been formulated as multi-stage stochastic programs (Govindan et al. 2017), and those that exist have tended to be small; e.g., a three-stage model with nine scenarios (Almansoori and Shah 2012).

2.4. Incentives and policy

The models in the previous subsections generally took the perspective of a company that aims to improve resiliency to reduce costs, though there is a stream of literature that considers incentives and policies from external decision-makers. Among those that included uncertainty in supply, researchers have analyzed strategies to incentivize capacity restoration after disruptions and to improve recovery time (Hu et al. 2013, Kim et al. 2010). As discussed, failure-to-supply clauses have been analyzed in the context of capacity allocation for drug shortages (Jia and Zhao 2017). Tang, Gurnani, and Gupta (2014) studied subsidies and increased demand to incentivize a more reliable supply. Researchers have also analyzed government policy incentives in other areas, e.g., tax incentives for renewable energy (Karimi et al. 2018).

2.5. Contributions

In this analysis of strategic resiliency decisions and policies to reduce drug shortages, this chapter make the following contributions.

- It studies resiliency decisions for the highly regulated supply chains of generic injectable drugs that have low profit margins and a limited ability to adapt if disruptions occur.
- I develop supply chain design models that incorporate a new combination of

characteristics. They consider time; disruptions may occur at multiple echelons; multiple components may be concurrently unavailable; and the company may select multiple mitigation strategies (facility redundancy and safety stock).

- It introduces constraints to enforce a replenishment rule, and the non-anticipativity property of the multi-stage stochastic program is induced. This allows a large thirteen-stage model to be solved.
- I evaluate policies to induce resiliency and reduce drug shortages (mandatory redundancy, mandatory inventory, failure-to-supply penalties, pricing changes, and the combination of price increases and other interventions). I analyze the social-efficiency of these interventions.

3. Base model (SCDD)

To begin to study strategic resiliency, I develop a two-stage stochastic program – the Supply Chain Design under Disruption (SCDD) model. In the first stage, the company selects the optimal configuration of the supply chain that will be fixed for the remainder of the time horizon. There is uncertainty about which components may be working in future periods. In each subsequent period, the uncertainty is realized, and the company decides the quantities of raw materials to order and finished goods to produce.

I consider a three-echelon supply chain for a single drug that is comprised of Active Pharmaceutical Ingredient (API) suppliers, manufacturing plants, and manufacturing lines. Manufacturers of sterile injectable drugs typically hold little safety stock (GAO 2016), and this is reflected in the base analysis.

3.1. Background

A sample supply chain configuration is presented in Figure 1. It includes two suppliers, two plants, and one line in each plant. Plants may receive raw materials from either API supplier, but lines are associated with specific plants. Each candidate plant $k \in K$ has a set of candidate lines $l \in L_k$. In this example, $L_1 = \{1,2,3\}$ and $L_2 = \{4,5,6\}$. The set of all lines L is the union of the sets of lines in each plant, i.e., $L = \cup_{k \in K} L_k = \{1,2,3,4,5,6\}$. While pharmaceutical companies have additional partners in practice (e.g., packaging and non-active raw materials), the model includes only the critical steps (cf. Bundschuh et al. 2003) and considers all echelons that contributed to shortage categories in a recent Government Accountability Office report (GAO 2016).

The objective is to maximize the expected profit under uncertainty in the status of the supply chain components. In the first stage, the company selects the supply chain configuration. They may choose to not market the drug and select no components. In the second stage, in each period, the company selects production and order quantities after uncertainty about component availability is realized. Demand may be met with production or unmet (a shortage), and both demand and price are constant over the time horizon.

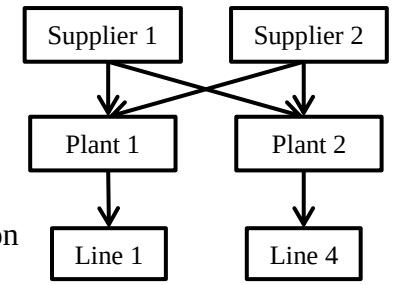

Figure 1. Example supply chain

There are fixed costs to select components, and these include the costs to maintain suppliers, plants, and lines as well as the government-mandated user fees for generic drug production (i.e., Generic Drug User Fee Amendments (GDUFA) fees; FDA 2018b). The GDUFA facility fees are incurred as fixed costs for each supplier and plant, and the GDUFA program fee is incurred if the company is in the market. There are variable costs to order raw materials and to produce the drug, and revenues come from sales.

Any component in any echelon may become disrupted, and in each period, the status of each component is either available $\{1\}$ or disrupted $\{0\}$. I model this uncertainty using discrete scenarios $\omega \in \Omega$, where the random variable ξ_{nt}^{ω} in scenario $\omega \in \Omega$ represents the status of candidate component $n \in N$ in period $t \in T$ ($\xi_{nt}^{\omega} \in \{0,1\}^{|N|x|T|x|\Omega|}$). The notation for SCDD is presented in Figure 2.

The lines have capacity limits to be consistent with the model that includes inventory, though these are not limiting for the SCDD model.

Sets

J	Set of candidate API suppliers
K	Set of candidate manufacturing plants
L	Set of candidate lines
L_k	Set of candidate lines in plant $k \in K$
N	Set of all components, $N = J \cup K \cup L$
T	Set of time periods
Ω	Set of scenarios

Decision Variables

First Stage

$x_j \coloneqq \begin{cases} 1 & \text{if API supplier } j \in J \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$

$y_k \coloneqq \begin{cases} 1 & \text{if manufacturing plant } k \in K \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$

$z_l \coloneqq \begin{cases} 1 & \text{if line } l \in L \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$

Second Stage

As a fraction of demand in period $t \in T$ in scenario $\omega \in \Omega$:

u_{jt}^{ω}	Raw material purchased from supplier $j \in J$
v_{lt}^{ω}	Finished goods produced on line $l \in L$
θ_t^{ω}	Demand met

Parameters

p^{ω}	Probability of scenario $\omega \in \Omega$
$\xi_{nt}^{\omega} \coloneqq \begin{cases} 1 & \text{if component } n \in N \text{ is available in period } t \in T \text{ in scenario } \omega \in \Omega \\ 0 & \text{otherwise} \end{cases}$	
d	Quantity of drug demanded each period
q	Sales price per unit of drug
c^{raw}, c^{prod}	Unit cost of raw materials and finished good production
$c^{API}, c^{Plant}, c^{Line}$	Annual fixed costs for each supplier, plant, and line, respectively
f^{API}, f^{Plant}	Annual GDUFA fees for each supplier and plant, respectively
$f^{Program}$	Annual GDUFA fee for drug program
$g^{Line} \in \mathbb{Z}^+$	Line capacity as a fraction of total demand, if line is available
$\ddot{t}$	Number of periods per year

Figure 2. Notation for SCDD

3.2. Model formulation

The formulation of SCDD is as follows.

Maximize

$$-\left\lceil\frac{|T|}{\ddot{t}}\right\rceil\left[(c^{API}+f^{API})\sum_{j\in J}x_j+(c^{Plant}+f^{Plant})\sum_{k\in K}y_k+c^{Line}\sum_{l\in L}z_l+f^{Program}x_1\right]+E_{\Omega}[Q(x,y,z)] \quad (1)$$

Subject to:

$$z_l \le y_k \qquad \forall k\in K, l\in L_k \quad (2)$$

$$x_j \ge x_{j+1} \qquad \forall j\in J\setminus\{|J|\} \quad (3a)$$
$$y_k \ge y_{k+1} \qquad \forall k\in K\setminus\{|K|\} \quad (3b)$$
$$z_l \ge z_{l+1} \qquad \forall l\in L_k\setminus\{|L_k|\}, k\in K \quad (3c)$$

$$x_j \in \{0,1\} \qquad \forall j\in J \quad (4a)$$
$$y_k \in \{0,1\} \qquad \forall k\in K \quad (4b)$$
$$z_l \in \{0,1\} \qquad \forall l\in L \quad (4c)$$

$$E_{\Omega}[Q(x,y,z)] = \max_{u,v,\theta}\sum_{\omega\in\Omega}p^{\omega}d\sum_{t\in T}\left[q\theta_t^{\omega}-c^{raw}\sum_{j\in J}u_{jt}^{\omega}-c^{prod}\sum_{l\in L}v_{lt}^{\omega}\right] \quad (5)$$

Subject to:

$$u_{jt}^{\omega} \le \xi_{jt}^{\omega}|L|g^{Line}x_j \qquad \forall j\in J, t\in T, \omega\in\Omega \quad (6)$$
$$v_{lt}^{\omega} \le \xi_{kt}^{\omega}\xi_{lt}^{\omega}g^{Line}z_l \qquad \forall l\in L_k, k\in K, t\in T, \omega\in\Omega \quad (7)$$
$$\sum_{l\in L}v_{lt}^{\omega} \le \sum_{j\in J}u_{jt}^{\omega} \qquad \forall t\in T, \omega\in\Omega \quad (8)$$
$$\theta_t^{\omega} \le \sum_{l\in L}v_{lt}^{\omega} \qquad \forall t\in T, \omega\in\Omega \quad (9)$$
$$\theta_t^{\omega} \le 1 \qquad \forall t\in T, \omega\in\Omega \quad (10)$$

$$u_{jt}^{\omega} \ge 0 \qquad \forall j\in J, t\in T, \omega\in\Omega \quad (11a)$$
$$v_{lt}^{\omega} \ge 0 \qquad \forall l\in L, t\in T, \omega\in\Omega \quad (11b)$$
$$\theta_t^{\omega} \ge 0 \qquad \forall t\in T, \omega\in\Omega \quad (11c)$$

The objective function (1) maximizes the expected profit. The annual fixed costs include, respectively, the cost per API supplier, the GDUFA fee per API supplier, the cost per plant, the GDUFA fee per plant, the cost per line, and the GDUFA program fee. Expected ordering and production costs and revenues are incurred in the second stage. Constraints (2) ensure that the selected lines are in selected plants. Constraints (3) require components to be selected in numerical order and are used to reduce alternative optima. Constraints (4) are

standard binary constraints.

In the second stage, the company makes operational decisions each period after uncertainty is realized. The objective function (5) maximizes the expected profit in the second stage. Revenues are accrued based on sales, and costs include raw materials and production. Constraints (6) limit orders of raw materials to selected, available suppliers. Constraints (7) limit finished goods production to the capacity of selected, available lines in available plants. Constraints (8) limit production to the amount of raw material ordered. Constraints (9-10) ensure the fraction of demand met is not greater than the finished goods available and the amount customers demand, respectively. Constraints (11) enforce non-negativity.

3.3. Structural property

It follows from the formulation of SCDD that demand is either fully met or fully unmet each period. This is presented formally in Lemma 1.

Lemma 1: $\theta_t^\omega \in \{0,1\}, \forall t \in T, \omega \in \Omega$. **Proof:** Provided in appendix.

3.4. Assumptions

To identify factors that contribute to supply chain vulnerability and resiliency, I make several simplifying assumptions. The model does not consider transportation time or cost. While these are clearly present in practice, the time to ship from common API supplier locations to the US is often one month or less (SeaRates 2018, US Department of Commerce 2018), which is smaller than the periods considered in this chapter's analyses. I assume production occurs throughout the year and suppliers are uncapacitated. I exclude some other operational dynamics such as costs and time of product changeover that are often present in other papers within the pharmaceutical literature (Lakhdar and Papageorgiou 2008, Marques et al. 2017). Discounting is not considered because the focus is on the realized costs and revenues of limited-term contracts.

If a disruption occurs, the model do not consider the cost of recovery. At an API facility, this cost would be incurred by the supplier, and at a plant the cost would be spread across multiple drugs, though this is a limiting assumption. I assume constant demand over the time horizon, consistent with the fairly stable demand of most drugs (Fox et al. 2014). If demand is not met for the drugs considered in our case examples, in practice a clinical decision is generally made to switch to an alternative treatment. Treatment delays are less common, and as a simplifying assumption, I assume all demand is lost rather than backordered due to delays. The model does not consider competition, which is justified by the fact that drugs affected by shortages are often sole-source, particularly those that are injectable (IMS Institute for Healthcare Informatics 2011, UUDIS 2016). The drugs analyzed have a single manufacturer.

I do not allow the company to make changes to the supply chain structure within the time horizon. Disruptions and recovery at each component occur independently and are exogenous to the model. Each candidate component within an echelon, e.g., all candidate lines, have identical disruption profiles and capacities. This could be relaxed by subscripting the capacity parameters by the components. Furthermore, following the work of others, I do not consider location decisions; rather the focus is on resiliency strategies and policies (Hopp and Yin 2006, Jia and Zhao 2017).

4. Extension to include inventory (SCDD-I)

To extend the analyses to consider safety stock as a resiliency strategy, I introduce a second model. It is called the Supply Chain Design model under Disruption with Inventory (SCDD-I) and formulate it as a multi-stage stochastic program. The model relaxes the initial decision that no inventory is held, and it can be used to study why companies may or may not use inventory.

4.1. Background

In the first stage, the company chooses the supply chain configuration and a target amount of safety stock to hold each period (i.e., stage). Neither the supply chain design nor the target safety stock level may be changed throughout the time horizon. As in the SCDD model, the company may choose no components and not produce the drug. In each of the subsequent stages, uncertainty in the component statuses is realized, and the company selects production and order quantities. Demand may be met through production or safety stock. If production exceeds demand, inventory is replenished. Unmet demand is lost, not backordered.

To model practice realistically, multi-stage stochastic programs must impose the non-anticipativity property. It requires decision-makers to make the same decision for each scenario that is identical up to that point. This prevents them from anticipating realizations of future uncertainty. The non-anticipativity property is typically enforced either through constraints or implied via the construction of the scenario tree, but as the number of stages increases, the number of constraints substantially increases, and the problem quickly becomes intractable to standard solution methods such as Sample Average Approximation (SAA). To avoid this

Decision Variables

First Stage

$\tilde{z}_{jl} \coloneqq \begin{cases} 1 & \text{if supplier } j \in J \text{ and line } l \in L \text{ are selected} \\ 0 & \text{otherwise} \end{cases}$

I_0 Target safety stock level as a fraction of per-period demand

Subsequent Stages

As a fraction of per-period demand in scenario $\omega \in \Omega$:

C_t^ω Capacity available to meet demand in period $t \in T$

$\tilde{C}_t^\omega$ Excess capacity available in period $t \in T$

I_t^ω Safety stock available at the end of period $t \in \{0\} \cup T$

In period $t \in T$ in scenario $\omega \in \Omega$:

$\delta_t^{Avail,\omega} \coloneqq \begin{cases} 1 & \text{if safety stock is available to meet demand} \\ 0 & \text{otherwise} \end{cases}$

$\delta_t^{Suffic,\omega} \coloneqq \begin{cases} 1 & \text{if sufficient capacity is available to replenish safety stock deficit} \\ 0 & \text{otherwise} \end{cases}$

Parameters

h Unit cost of holding drug for one period

o^{max} Maximum number of periods of safety stock that can be held

Figure 3. Additional Notation for SCDD-I

intractability, I impose a safety stock replenishment rule: the manufacturer must meet demand when possible; they may only deplete safety stock if production capacity is unavailable; and they must replenish deficit safety stock if excess capacity is available. That is, given the component statuses and variables from the previous stage, the rule predetermines the sales and inventory decisions; the decision variables will be the same regardless of future realizations of uncertainty. This induces the non-anticipativity constraints to hold without including them in the model. The additional notation for SCDD-I is presented in Figure 3.

4.2. Model formulation

The SCDD-I model includes constraints (2-4, 6-8, 10-11) from the SCDD model, and the revised objectives and additional constraints are as follows.

Maximize

$$-\left\lceil\frac{|T|}{\ddot{t}}\right\rceil\left[(c^{API}+f^{API})\sum_{j\in J}x_j+(c^{Plant}+f^{Plant})\sum_{k\in K}y_k+c^{Line}\sum_{l\in L}z_l+f^{Program}x_1\right]+E_{\Omega}[Q(x,y,z,I_0)] \quad (12)$$

Subject to:

$$I_0 \le o^{max}\tilde{z}_{11} \quad (13)$$

$$\tilde{z}_{jl} \ge x_j + z_l - 1 \qquad \forall j\in J, l\in L \quad (14a)$$
$$\tilde{z}_{jl} \le x_j \qquad \forall j\in J, l\in L \quad (14b)$$
$$\tilde{z}_{jl} \le z_l \qquad \forall j\in J, l\in L \quad (14c)$$

$$I_0 \ge 0 \quad (4d)$$
$$\tilde{z}_{jl} \in \{0,1\} \qquad \forall j\in J, l\in L \quad (4e)$$

$$E_{\Omega}[Q(x,y,z,I_0)] = \max_{u,v,\theta,I}\sum_{\omega\in\Omega}p^{\omega}d\sum_{t\in T}\left[q\theta_t^{\omega}-c^{raw}\sum_{j\in J}u_{jt}^{\omega}-c^{prod}\sum_{l\in L}v_{lt}^{\omega}-hI_t^{\omega}\right] \quad (15)$$

Subject to:

$$\theta_t^{\omega} \le \sum_{l\in L}v_{lt}^{\omega}+I_{t-1}^{\omega} \qquad \forall t\in T,\omega\in\Omega \quad (16)$$
$$I_t^{\omega} = I_{t-1}^{\omega}+\sum_{l\in L}v_{lt}^{\omega}-\theta_t^{\omega} \qquad \forall t\in T,\omega\in\Omega \quad (17)$$
$$I_t^{\omega} \le I_0 \qquad \forall t\in T,\omega\in\Omega \quad (18)$$

$$\theta_t^{\omega} \ge C_t^{\omega} \qquad \forall t\in T,\omega\in\Omega \quad (19)$$

$$\theta_t^\omega \geq \delta_t^{Avail,\omega} \qquad \forall t \in T, \omega \in \Omega \qquad (20)$$
$$I_{t-1}^\omega - I_t^\omega \leq 1 - C_t^\omega \qquad \forall t \in T, \omega \in \Omega \qquad (21)$$

$$C_t^\omega \geq \xi_{jt}^\omega \xi_{kt}^\omega \xi_{lt}^\omega \tilde{z}_{jl} \qquad \forall t \in T, \omega \in \Omega, j \in J, k \in K, l \in L_k \qquad (22a)$$
$$C_t^\omega \leq 1 \qquad \forall t \in T, \omega \in \Omega \qquad (22b)$$
$$C_t^\omega \leq \sum_{j\in J} \sum_{k\in K} \sum_{l\in L_k} \xi_{jt}^\omega \xi_{kt}^\omega \xi_{lt}^\omega \tilde{z}_{jl} \qquad \forall t \in T, \omega \in \Omega \qquad (22c)$$
$$\tilde{C}_t^\omega = \sum_{j\in J} \sum_{k\in K} \sum_{l\in L_k} g^{Line} \xi_{jt}^\omega \xi_{kt}^\omega \xi_{lt}^\omega \tilde{z}_{jl} - C_t^\omega \qquad \forall t \in T, \omega \in \Omega \qquad (23)$$

$$\sum_{l\in L} v_{lt}^\omega - C_t^\omega = \min\left(\tilde{C}_t^\omega, I_0 - I_{t-1}^\omega\right) \qquad \forall t \in T, \omega \in \Omega \qquad (24\text{-nonlin})$$
$$\sum_{l\in L} v_{lt}^\omega - C_t^\omega \geq \tilde{C}_t^\omega - |L| g^{Line} \delta_t^{Suffic,\omega} \qquad \forall t \in T, \omega \in \Omega \qquad (24a)$$
$$\sum_{l\in L} v_{lt}^\omega - C_t^\omega \geq I_0 - I_{t-1}^\omega - (o^{max} + \epsilon)\left(1 - \delta_t^{Suffic,\omega}\right) \quad \forall t \in T, \omega \in \Omega \qquad (24b)$$

$$1 - \delta_t^{Avail,\omega} \geq 1 - I_{t-1}^\omega \qquad \forall t \in T, \omega \in \Omega \qquad (25a)$$
$$(o^{max} + \epsilon)\delta_t^{Avail,\omega} \geq I_{t-1}^\omega \qquad \forall t \in T, \omega \in \Omega \qquad (25b)$$

$$I_0^\omega = I_0 \qquad \forall \omega \in \Omega \qquad (26)$$

$$I_t^\omega \geq 0 \qquad \forall t \in \{0\} \cup T, \omega \in \Omega \qquad (27a)$$
$$C_t^\omega, \tilde{C}_t^\omega \geq 0 \qquad \forall t \in T, \omega \in \Omega \qquad (27b)$$
$$\delta_t^{Suffic,\omega}, \delta_t^{Avail,\omega} \in \{0,1\} \qquad \forall t \in T, \omega \in \Omega \qquad (27c)$$

The objective function (12) is the same as (1) except it adds the target inventory as an argument to the expected cost of the subsequent stages. In the first stage, constraint (13) enforces two safety stock conditions. It requires a complete supply chain to be selected if safety stock is held, via the binary variable, z_{11}, and it limits the number of periods of safety stock that can be held to an upper bound, o^{max}. Constraints (14) define variables to indicate which combinations of suppliers and lines are selected. These are used in combination with constraints (22-23) to define the total capacity available in each period. Constraints (4d-e) enforce the domains of the new first stage decision variables.

In the subsequent stages, the objective function (15) maximizes expected profit. Revenues come from demand met. The variable costs are incurred through raw material orders; production of finished goods; and held safety stock, respectively. Constraints (16) ensure that demand can only be met from production and safety stock. Constraints (17) provide safety stock

balance across the time periods. The safety stock remaining at the end of a period is equal to the amount held-over from the previous period plus the finished goods produced minus the amount used to meet demand. Constraints (18) prevent the manufacturer from holding more safety stock than the selected target level.

Constraints (19-20) enforce the rule that the company must meet demand, if possible. Constraints (19) require demand to be satisfied when there is production capacity, and constraints (20) require demand to be met when there is safety stock. Note that the capacity-to-meet-demand variable C_t^{ω} is implied to be binary, proven via Lemma 2 in Section 4.3. Constraints (21) only allow safety stock to be depleted if there is no available capacity.

Constraints (22-23) define the two capacity-related variables: capacity-to-meet-demand and excess capacity, respectively. Constraints (22a-b) ensure the capacity-to-meet demand variables are 1 if a complete supply chain is working and selected, and constraints (22c) require them to be 0, if not. More specifically, in constraints (22a), the coefficient $\xi_{jt}^{\omega}\xi_{kt}^{\omega}\xi_{lt}^{\omega}$ represents the status (available or disrupted) of a complete supply chain (j, k, l) of a given supplier $j \in J$ and line $l \in L_k$ in plant $k \in K$ in period $t \in T$. For a complete supply chain to be available, each component in the configuration must be available, i.e., $\xi_{jt}^{\omega} = \xi_{kt}^{\omega} = \xi_{lt}^{\omega} = 1$. The variable $\tilde{z}_{jl}$ designates whether the complete supply chain that includes supplier $j \in J$ and line $l \in L$ is selected. If there is a complete supply chain selected and available, the right-hand side of constraint (22a) will be 1 for at least one combination, and the capacity-to-meet demand variable will be forced to at least 1; it is limited to 1 via constraints (22b). In constraints (22c), the term $\sum_{j\in J}\sum_{k\in K}\sum_{l\in L_k}\xi_{jt}^{\omega}\xi_{kt}^{\omega}\xi_{lt}^{\omega}\tilde{z}_{jl}$ sums the statuses of candidate complete supply chains, and if there is not a selected and available supply chain, the capacity-to-meet demand variable is forced to 0. Constraints (23) define the excess capacity available each period. It is calculated as the total

available capacity minus the capacity-to-meet demand.

Constraints (24-nonlin) enforce the requirement that the company must replenish safety stock when possible, up to the target level. As these are nonlinear, I reformulate them using binary indicator variables and implement constraints (24a-b) rather than constraints (24-nonlin). The values of the left-hand sides of constraints (24a-b) represent the amount of drug produced over the amount used to meet demand. The constraints require that these values must be at least equal to the excess capacity (24a) or the safety stock deficit (24b).

Constraints (25) indicate whether safety stock is available to meet demand. Constraints (25a) force the availability indicator to 0 if no safety stock is held-over from the previous period, and constraints (25b) force it to 1 if there is. Note that $I_t^\omega \ \forall t \in \{0\} \cup T, \omega \in \Omega$ will not take on fractional values, as stated by Lemma 4 in Section 4.3. Constraints (26) set the safety stock levels at the beginning of the time horizon to the selected target level. Constraints (27) are standard non-negativity and domain constraints.

4.3. Structural properties

In this section, I present key structural properties of the SCDD-I model. In Lemma 2, the capacity-to-meet demand variables are implied to be binary, and Lemma 3 states that the excess capacity variables are implied to be integer. Lemma 4 indicates that the number of periods of inventory held are implied to be integer. Lemma 5 states that demand is either fully met or fully unmet each period.

Lemma 2: $C_t^\omega \in \{0,1\}, \forall t \in T, \omega \in \Omega$. **Proof:** Provided in appendix.

Lemma 3: $\tilde{C}_t^\omega \in \mathbb{Z}^+, \forall t \in T, \omega \in \Omega$. **Proof:** Provided in appendix.

Lemma 4: $I_0, I_t^\omega \in \mathbb{Z}^+, \forall t \in \{0\} \cup T, \omega \in \Omega$. **Proof:** Provided in appendix.

Lemma 5: $\theta_t^\omega \in \{0,1\}, \forall t \in T, \omega \in \Omega$. **Proof:** Provided in appendix.

I further use these lemmas and the corresponding safety stock replenishment rule to establish the implied non-anticipativity of SCDD-I through Theorem 1. The manufacturer's decisions are only based on variables for the previous stage and the realization of uncertainty at the current stage; they do not consider uncertainty that will subsequently be revealed. For each period, it is optimal to make the same decisions for each of the scenarios that have identical realizations of uncertainty up to that period. I define S_t^ω as the set of scenarios that have paths that are indistinguishable from scenario $\omega \in \Omega$ in period $t \in \{0\} \cup T$.

Theorem 1: The following relationships are implied by SCDD-I.

$$C_t^\omega = C_t^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28a)$$
$$\tilde{C}_t^\omega = \tilde{C}_t^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28b)$$
$$I_t^\omega = I_t^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28c)$$
$$\theta_t^\omega = \theta_t^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28d)$$
$$\sum_{l \in L} v_{lt}^\omega = \sum_{l \in L} v_{lt}^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28e)$$
$$\sum_{j \in J} u_{jt}^\omega = \sum_{j \in J} u_{jt}^{\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28f)$$
$$\delta_t^{Avail,\omega} = \delta_t^{Avail,\omega'} \qquad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28g)$$
$$\delta_t^{Suffic,\omega} = \delta_t^{Suffic,\omega'} \text{ except case: } \tilde{C}_t^\omega = I_0 - I_{t-1}^\omega \ \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \qquad (28h)$$

Proof: Provided in appendix.

4.4. Additional assumptions

SCDD-I is subject to the non-inventory-related assumptions of SCDD that are discussed in Section 3.4. In addition, because both the raw materials and finished form of the drugs are perishable, I make two assumptions: the company may not hold raw material inventory, and there is an exogenous limit to the amount of finished goods inventory that may be held. In conversations with a pharmaceutical manufacturer, these are consistent with practice. The model assumes that inventory is not destroyed if a facility is disrupted. Finally, the capacity of each line as a fraction of per-period demand is required to be integer-valued and at least equal to 1, and I evaluate different values in scenario analyses.

5. Solution methods

The problems presented in Sections 3 and 4 represent a two- and a multi-stage stochastic program, respectively. There are $2^{|N|}$ combinations of possible statuses for the candidate components, and over the entire time horizon, this produces a full scenario set of $\left(2^{|N|}\right)^{|T|}$. For 10 candidate components and 4 time periods, this produces a set of $1.1 * 10^{12}$ scenarios and for 12 time periods produces a set of $1.3 * 10^{36}$ scenarios. As this is large, I approximate the optimal value using the SAA algorithm (Kleywegt et al. 2002).

The implementation is presented in Figure 4. The algorithm is comprised of three key steps – optimization, solution evaluation, and bound calculation. *SCDD*$_{SAA}$ represents either the SCDD or SCDD-I model, as appropriate, where the complete uncertainty set Ω is replaced by a set of sampled scenarios. In the optimization step, the set of τ sampled scenarios is $\widetilde{\Omega}^r, \forall r \in \{1, \ldots, R\}$. In the evaluation step, the set of $\check{\tau}$ sampled scenarios is $\check{\Omega}$. The optimal value of each solution evaluation step, $\check{V}^r, \forall r = \{1, \ldots, R\}$, is an approximate lower bound on the true optimal value, V^*. To set the lower bound, I select r' to be the index of the median value of $\check{V}^r$.

1. Optimization and approximation
   a. For $r = 1, \ldots, R$, sample τ independently and identically distributed scenarios ξ_{nt}^{ω} from the complete uncertainty set Ω
      i. Define the set of these scenarios to be $\widetilde{\Omega}^r$
      ii. Solve *SCDD*$_{SAA}$ with uncertainty set $\widetilde{\Omega}^r$
      iii. Record the optimal value, V^r, and the first stage decision variables, x^r, y^r, z^r, I_0^r
2. Evaluation
   a. Generate $\check{\tau}$ independently and identically distributed samples of ξ_{nt}^{ω} from the complete uncertainty set Ω
      i. Define the set of these scenarios to be $\check{\Omega}$
      ii. For $r = 1..R$,
         1. Fix $x = x^r, y = y^r, z = z^r, I_0 = I_0^r$
         2. Solve *SCDD*$_{SAA}$ with uncertainty set $\check{\Omega}$
         3. Record the optimal value, $\check{V}^r$
3. Computation of the lower and upper bounds
   a. Select replication index r', and set $LB := \check{V}^{r'}$
   b. Set $UB := \frac{1}{R}\sum_{r=1}^{R} V^r$

Figure 4. Implementation of SAA

6. Case study: two generic oncology drugs

To evaluate potential policy effects, two drugs are considered as case examples – vinblastine sulfate and vincristine sulfate. Both are generic, injectable drugs produced by single manufacturers that have been subject to recent shortages in the US (UUDIS 2016). Vinblastine is used to treat various cancers including testicular cancer and lymphomas, and vincristine is used to treat leukemias and lymphomas (Drugs.com 2018). Both are curative for some conditions, and they were selected based on conversations with an oncology pharmacist and a review of the literature.

6.1. Data

Table 2 presents data that are used for both analyses. Table 3 presents data that are specific to each drug, including demand and costs. These are derived from the available literature and conversations with subject matter experts.

The total US demand of each drug is estimated based on Medicare Part B data (CMS 2018a, b), and prices are applied from the Red Book (IBM Micromedex 2018). Raw material

costs are estimated from conversations with suppliers and available data online (PharmaCompass 2018). Some of the costs are proprietary, but conversations with an industry expert estimated the full cost to produce a drug are 20-60% of the drug price, consistent with the values used by Jia and Zhao (2017). Using this range and other cost values, I calculated the production costs for each drug and non-fee-related fixed costs. Details are included in the supplementary materials, and these values were tested in sensitivity analyses. Based on conversations with an industry expert, the analyses allow up to 2 years of finished goods inventory to be held, and capacities for the suppliers and lines are assumed.

The distributions of time to disruption are estimated based on FDA data on drug approval dates and the start dates of shortages reported by University of Utah Drug Information Service (UUDIS) (FDA 2018a, UUDIS 2016). The distributions of time to recover are estimated from UUDIS on shortage length (UUDIS 2016). Based on the data, geometric distributions are applied for both disruption and recovery, and further detail is available in the supplementary material. I consider a two-year time horizon based on conversations with the procurement office at a large academic health system and apply two-month time periods to be sufficiently granular while maintaining feasible run times. For the SCDD model, I apply 30 replications (R), 600 optimization scenarios (τ), and 1,200 evaluation scenarios ($\check{\tau}$). For SCDD-I, I apply 40, 100, and 1,500, respectively. These were calibrated to consistently produce optimality gaps of 1% (SCDD) and 2% (SCDD-I). When the parameter values are at their baseline values, the SCDD model required approximately 340 seconds to run, and the SCDD-I required approximately 3,000 seconds.

Table 2. Values for all analyses[§]

Input	Supplier	Plant	Line	Source
Annual fixed costs	$33,000	$65,000	$32,500	Rudge (2012) and assumptions
Annual GDUFA fees	$1,169	$4,401	n/a	Calculated based on FDA (2018b)
Capacity as a fraction of per-period demand	n/a	n/a	2	Assumed
Average time to disruption, in years	17.3	28.2	8.5	Calculated based on FDA (2018a) and UUDIS (2016)
Average time to recovery, in years	1.2	0.8	0.08	Calculated from UUDIS (2016)

API = Active Pharmaceutical Ingredient; GDUFA = Generic Drug User Fee Amendments; SAA = Sample Average Approximation
[§]Costs in 2018 US dollars

Table 3. Drug-specific data[§]

Input	Vinblastine	Vincristine	Source
Annual demand, in ml	315,000	90,000	Estimated from CMS (2018a, b) and National Cancer Institute (2018)
Price, per ml	$4.31	$5.55	WAC reported by IBM Micromedex (2018); Vincristine sulfate price taken as average of WAC per unit.
Raw material cost per ml	$0.23	$0.34	Procurement representatives; PharmaCompass (2018)
Production cost per ml	$1.16	$2.22	Calculated
Annual holding cost per ml	$1.55	$2.00	Based on 3% monthly from Jia and Zhao (2017)
Annual GDUFA program fee	$11,445	$9,700	Calculated based on FDA (2018b, c)

ml = milliliter; WAC = Wholesale Acquisition Cost
[§]Costs in 2018 US dollars

6.2. Analysis results

Using the two models, I analyze how companies design their supply chains under different conditions and the associated impact on shortages and profit. All of the analyses were conducted with both the SCDD and SCDD-I versions of the model, except for the safety stock analysis. When inventory is not selected in the optimal solution for either drug, the SCDD are presented.

For each policy analysis, I present figures with 3 or 4 panels for each drug. These are: the optimal supply chain configuration; the target number of periods of inventory to hold (if selected); the expected shortage, and the percent difference in profit versus baseline. In some cases, there is apparent variability within a given cluster of points; this is largely due to the fact that the SAA method does not guarantee exact optimality. In the text, I round the values for unmet demand to the nearest percent and profit to the nearest $1,000. The algorithm and model were programmed in AMPL and solved using CPLEX 12.7 (Fourer et al. 2002, IBM 2017). The analyses were conducted on a PC with a 2.3 GHz Intel Core i7 and 16 GB of RAM.

No Intervention (Baseline)

In the base-case, no policies are imposed. The manufacturer of vinblastine selects 2 suppliers, 1 plant, 1 line, and no safety stock. The expected percent of demand that is not satisfied (shortage) is 6% with a corresponding expected annual profit of $686,000. For vincristine (a higher cost, lower demand drug), the manufacturer selects 1 supplier, 1 plant, and 1 line with no safety stock, and the expected annual profit is $93,000. The expected shortage is 11%.

Redundancy

One proposal to increase resiliency is to require a company to maintain multiple components at a single echelon. This has been noted in the FDA Strategic plan (FDA 2013), a

report from the Drug Shortages Summit (ASHP 2013), a joint letter to Congress from major health organizations (AHA et al. 2017), and other literature (Chabner 2011, Gehrett 2012, Health Policy Brief: Drug Shortages 2014, Jarosławski et al. 2017). To test the effects of redundancy regulation, I add the following variables and constraints.

New decision variables

$$\tilde{\delta}_k \coloneqq \begin{cases} 1 & \text{if at least one line in plant } k \in K \text{ is selected} \\ 0 & \text{otherwise} \end{cases}$$

Constraints

$$\sum_{j \in J} x_j \geq 2x_1 \tag{29}$$
$$\sum_{k \in K} y_k \geq 2y_1 \tag{30}$$
$$\sum_{l \in L} z_l \geq 2z_1 \tag{31}$$

$$\tilde{\delta}_k - z_l \geq 0 \qquad \forall l \in L_k, k \in K \tag{32}$$
$$\tilde{\delta}_k \leq \sum_{l \in L_k} z_l \qquad \forall k \in K \tag{33}$$
$$\sum_{k \in K} \tilde{\delta}_k \geq 2y_1 \tag{34}$$

These constraints mandate that the company have multiple components at the given echelon(s) if they choose to be in the market. Constraints (29-31) require two suppliers, two plants, and two lines to be selected given that one component is selected, respectively. Constraints (32-34) require lines to be selected in multiple plants. Constraints (32-33) assign variables to indicate whether a line is selected in each plant. Constraint (34) requires lines to be selected in at least two plants if any plants are selected. To require the company to have multiple API suppliers, I include constraint (29). To require multiple plants, constraints (32-34) are included, and to require multiple lines, constraint (31) is included. To enforce redundancy at all levels, I include constraints (29-31).

The results for this analysis are presented in Figure 5. For both drugs across all regulations, the company selects no inventory and adds exactly as much redundancy as is required, unless it is unprofitable. For vinblastine, the company continues to maintain a second

supplier, even when it is not mandatory, as in the baseline analysis. Any level of required redundancy is profitable for vinblastine, and the resiliency decisions lead to shortages of 1-6%, varying by echelon. Redundancy at all levels reduces the shortage to 1% but is the most costly; the expected annual profit decreases by 8%.

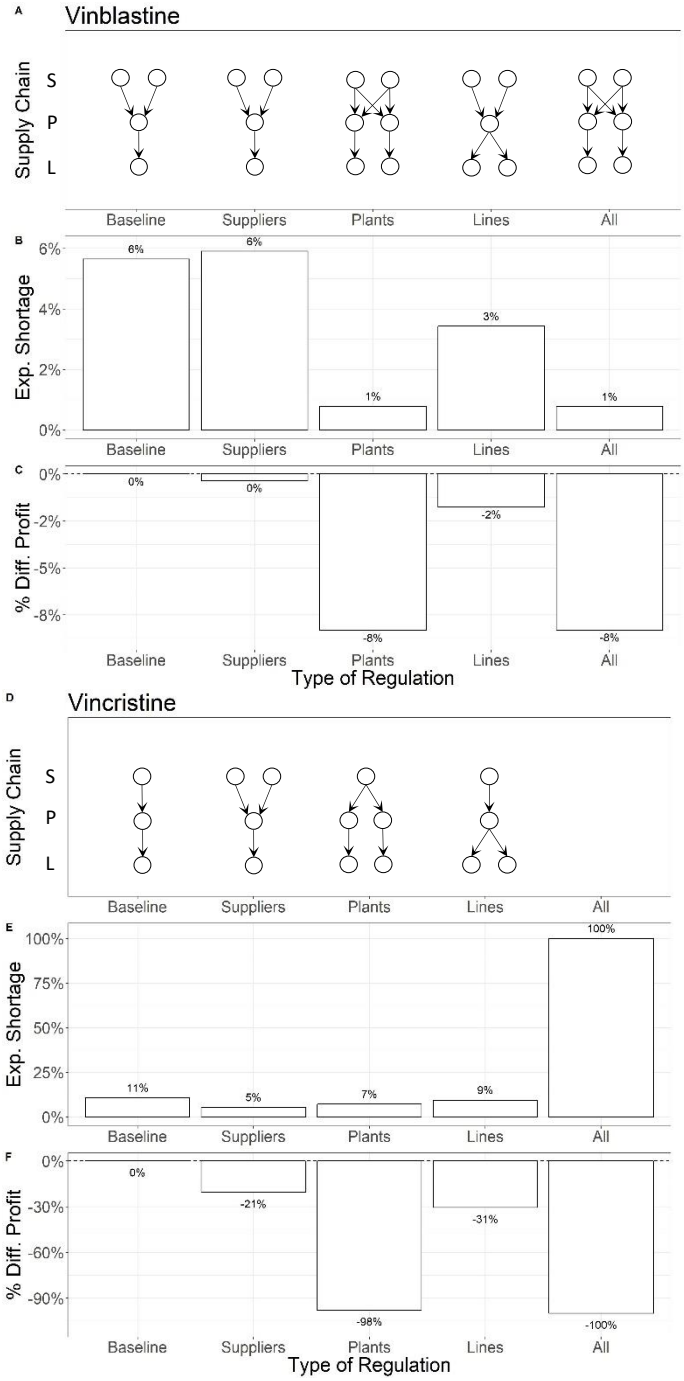

Figure 5. Effects of redundancy regulations

For vincristine, when redundancy at a single level is mandatory, the expected shortage drops to 5-9% of demand. Shortages are lowest (5% of demand) when a backup supplier is required, and expected profit is 21% lower than at baseline. If redundancy is required at every echelon, it would be unprofitable to make the drug, and the company chooses not to produce it (an expected shortage of 100%).

For both drugs, requiring a second plant causes substantial declines in expected profit (8% for vinblastine; 98% for vincristine). The costs and fees to maintain an additional plant and line are high relative to the baseline profits, and the increases in revenue from providing more of the drugs do not fully cover them. In general, redundancy regulations affect the difference in expected profit of vincristine more than vinblastine; this occurs because the baseline profit of vincristine is substantially lower. For a given regulation, the expected shortages for vincristine are generally higher than vinblastine because the vincristine supply chain does not include a backup supplier unless mandated.

Mandatory Inventory Levels

Some have proposed requiring manufacturers to hold minimum levels of inventory (e.g., ASHP 2013, FDA 2013, Gupta and Huang 2013). To run these analyses, I add a new parameter, $\widetilde{\Pi}$, to represent the minimum level of target safety stock if the manufacturer is in the market. I also add constraint (35) to require that the target safety stock level be at least the minimum if any plants are selected. These analyses were run using the SCDD-I model.

$$I_0 \geq \widetilde{\Pi} y_1 \qquad (35)$$

For both drugs, the manufacturer holds exactly the amount of inventory required up to a threshold at which it becomes unprofitable (Figure 6). For vinblastine, the company holds up to 20 months of inventory and for vincristine, up to 8 months. When at least 4 months are required,

the vinblastine manufacturer does not maintain a backup supplier and uses inventory as the sole resiliency strategy. At 6 months of inventory, the expected shortages of vinblastine and

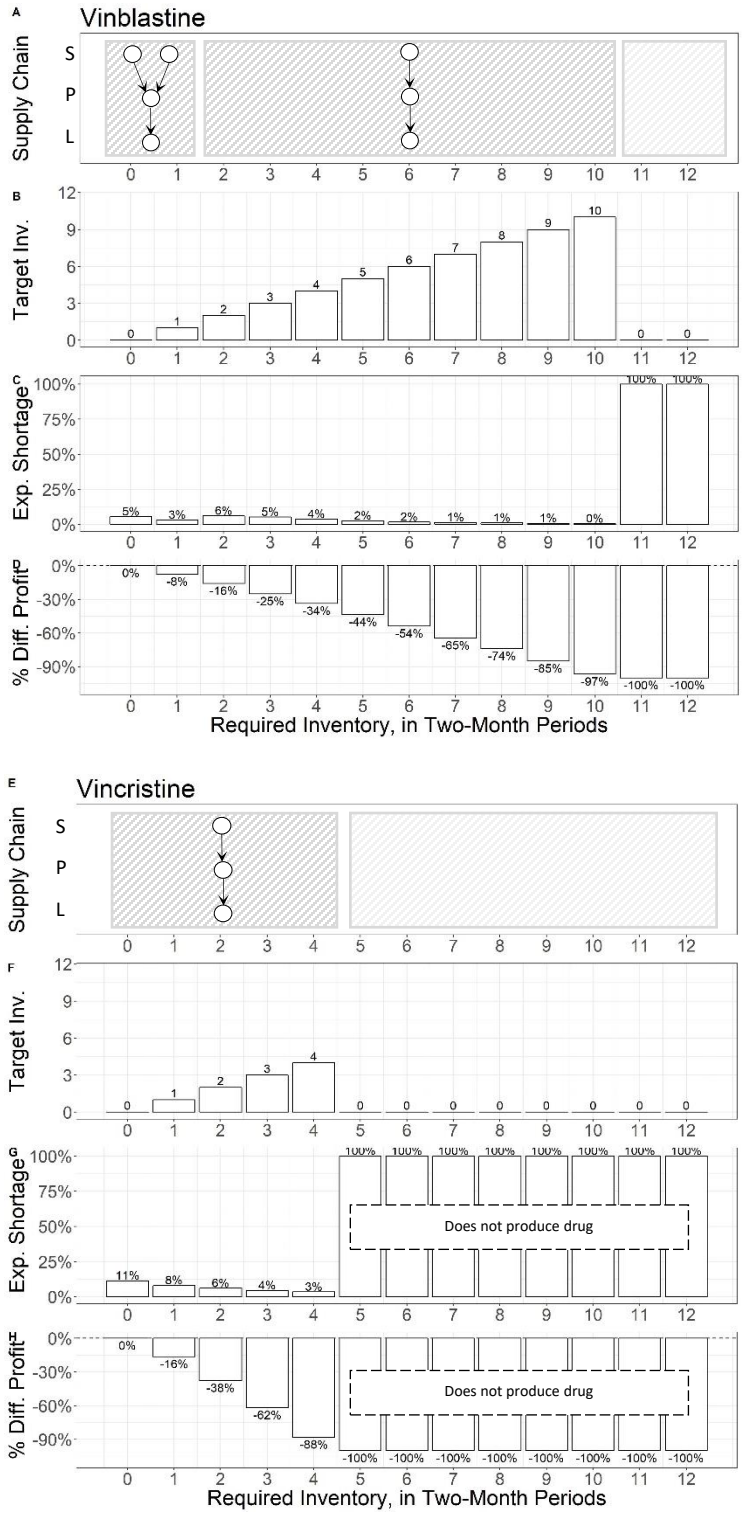

Figure 6. Effects of safety stock requirements

vincristine are 5 and 4%, respectively, with drops in profit vs. baseline of 25% and 62%.

As the amount of mandatory inventory increases, the expected profits decrease. The expected shortages also generally decrease, up until the point where the drugs are not produced. The one exception occurs for vinblastine when the inventory requirement is increased from 2 months to 4 months. At 4 months of inventory, the company no longer maintains a backup supplier (Figure 6; panel A). This occurs because they are optimizing for expected profit, rather than for shortages, and with four months of inventory, it is more profitable to maintain a single supplier than multiple (even though 4 and 6 months of inventory provide less protection against expected shortages than a backup supplier and 2 months of inventory).

Failure-to-Supply Penalties

Pharmaceutical contracts typically do not include strong penalties if the manufacturer cannot supply the drug. If penalties are included, contracts are often written to require reimbursement for the additional cost to purchase the same drug from a different manufacturer (Haninger et al. 2011, Jia and Zhao 2017). However, frequently the drug is not available elsewhere, and these penalties are rarely paid. Several researchers have suggested that strengthening failure-to-supply clauses may induce resiliency (Conti 2011, FDA 2013, Haninger et al. 2011, Health Policy Brief: Drug Shortages 2014, Jia and Zhao 2017, Reed et al. 2016). In this analysis, I apply a failure-to-supply penalty for each unit of unmet demand and add the term $-c^{short}(1-\theta_t^{\omega})$ to the objective functions (5) and (15). I present results for the SCDD-I model (Figure 7).

As it becomes more costly to not meet demand, the companies choose to add resiliency, and the expected shortages decrease. Resiliency is added at thresholds of the failure-to-supply penalties. Between these thresholds, as the penalty values increase, there is no change in the

resiliency decisions, nor by extension, in the expected shortages. For example, when the failure-to-supply penalty for vincristine is $3.89 (70% of price), the company adds a backup supplier, and the shortages decrease from 12% to 5%. At the next threshold, $11.10 (200% of price), the company chooses to hold 2 months of safety stock and shortages drop to 2%. Between these thresholds, as the penalty is increased, the expected profit declines though the expected shortages stay fairly consistent (Figure 7, panels G and H). It becomes unprofitable to produce vincristine when the penalty is at least $22.20 (400% of price).

The thresholds at which failure-to-supply penalties change the resiliency decisions for vinblastine are $2.16 (50% of the unit price) when the company adds a backup line, and at $6.47 (150% of price) when the company adds a backup plant. These additions reduce the expected shortages to 3% and 1%, respectively.

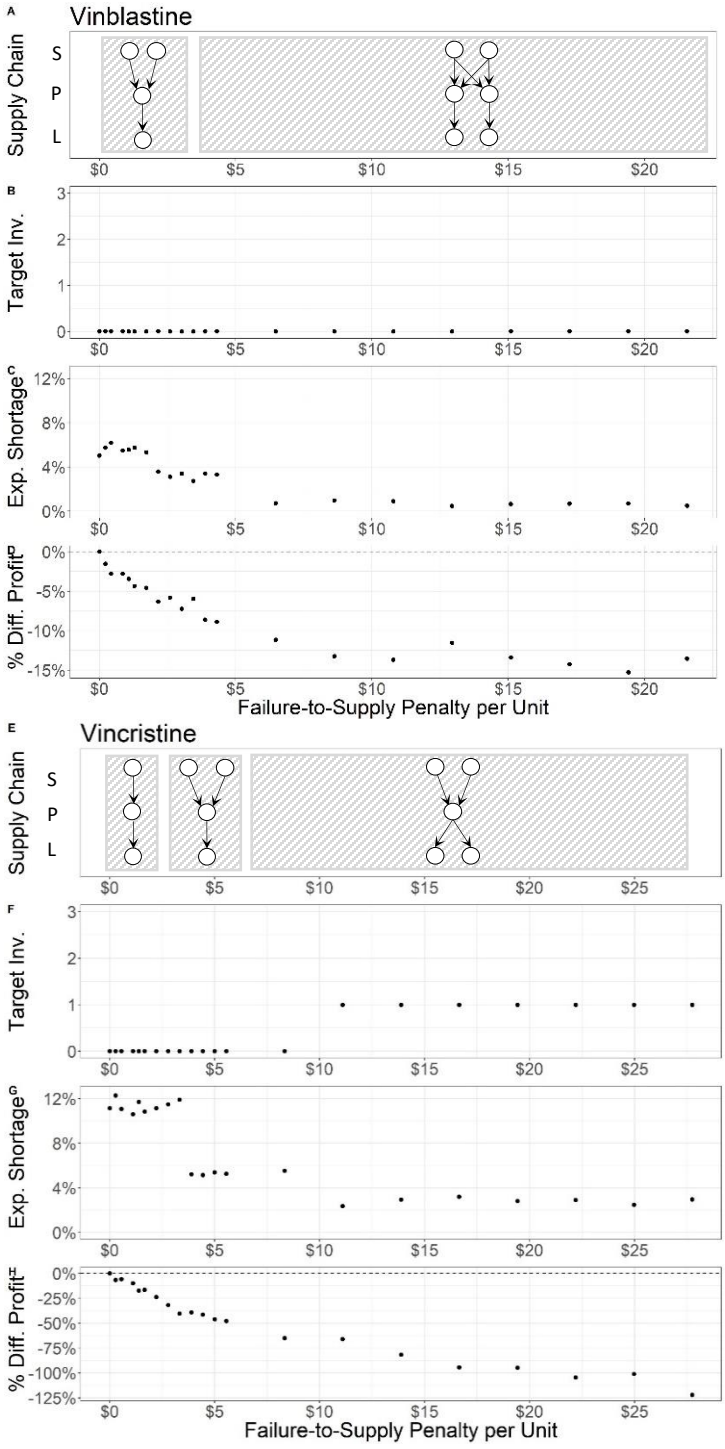

Figure 7. Effects of failure-to-supply penalties

Pricing

Some experts have pointed to the low prices of certain types of drugs as a primary driver of drug shortages. There have been corresponding calls for higher prices (e.g., Chabner 2011; Frakt 2016; Gatesman and Smith 2011; "Health Policy Brief: Drug Shortages," 2014; Link et al. 2012). In this analysis, I vary the prices of each drug (Figure 8).

As the prices increase, the manufacturers of vinblastine and vincristine add more resiliency and expected shortages decline. As the price increases, the opportunity cost for not providing the drug during periods of shortage increases; at certain thresholds, it becomes more profitable to invest in resiliency and to be able to provide the drug more often. For vinblastine, when the price is 2 times baseline, the company adds a second line, and at 2.5 times baseline, it adds a second plant; the expected shortages are 3% and 1%, respectively. The corresponding expected profits increase 189% and 286% vs. baseline. For prices between 2.5 and 10 times baseline, the company does not add resiliency though expected profit continues to increase nearly linearly.

For vincristine, the price thresholds at which the company changes its supply chain are 1.75 times baseline (adds backup supplier); 2.5 times baseline (adds 2 months of inventory); and 9 times baseline (adds backup plant and removes inventory). The corresponding expected shortages are 5%, 3%, and 1%, and the difference in expected profits are 344%, 759%, and 4,214% vs. baseline, respectively.

The model can also be used to analyze the potential effects of price declines. If the price of vinblastine drops to 70% of baseline, the company does not maintain a second supplier, and the expected shortage increases from 5% to 11%. For vincristine, if the price drops to 75% of baseline, the company does not produce the drug.

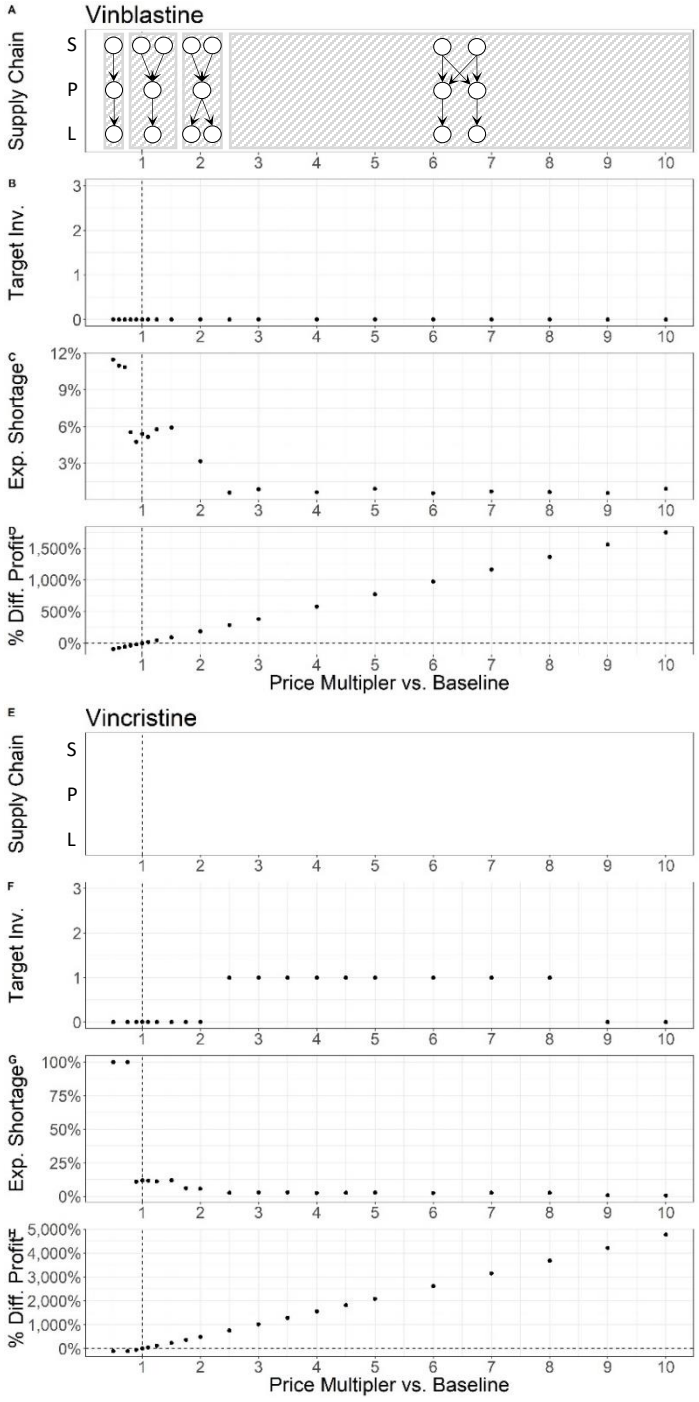

Figure 8. Effects of varying drug price

Social-Efficiency

Many of the proposed policies increase cost, and prices could be concurrently increased to mitigate the effects on company profit. In this section, I analyze how much prices would need to increase to maintain expected profits at approximately baseline levels and calculate the societal costs to achieve target shortage levels. Table 4 presents the policies that would lead to expected shortages of at most 5% and 2%, meaning the drug is available 95% and 98% of the time, respectively. The societal costs of each policy are calculated as the extra amount paid annually due to price increases, i.e., the product of the baseline drug price, annual demand, and the percentage price increase. For each policy analysis, I incrementally increased the price until the expected profit was approximately the baseline level.

For vinblastine, shortages at baseline are approximately 5%, and no intervention is necessary to reach this threshold. To achieve expected shortages of at most 2%, the following policies in combination with price increases would be effective: requiring a backup plant, failure-to-supply penalties of 150% of price, or 12 months of inventory; a price increase of 150% without other intervention would also be effective. The societal costs of each are $136,000 to $2 million, varying by policy.

To achieve expected vincristine shortages of at most 5%, requiring a backup supplier and increasing the price by 10% has the lowest societal cost. For expected shortages of 2%, price increases of 30% in combination with requiring redundancy at all levels or 12 months of inventory have the lowest societal costs.

Table 4. Summary of policy costs

Drug	Shortage Upper Bound	Policy	Price Increase	Annual Societal Cost
Vinblastine	0.05	Not applicable[§]	0%	$0
	0.02	Require multiple plants	10%	$136,000
		150% failure-to-supply penalty	10%	$136,000
		Require 12 months inventory	30%	$407,000
		Price increase	150%	$2,036,000
Vincristine	0.05	Require multiple suppliers	10%	$50,000
		70% failure-to-supply penalty	20%	$100,000
		Require 6 months of inventory	20%	$100,000
		Price increase	70%	$350,000
	0.02	Require multiple redundancy at all levels	30%	$150,000
		Require 12 months of inventory	30%	$150,000
		200% failure-to-supply penalty	50%	$250,000
		Price increase	800%	$3,996,000

[§]The results at baseline approximately achieve the 5% threshold (using the SCDD model: 6%; SCDD-I model: 5%).

6.3. Sensitivity, scenario, and validation analyses

To analyze the sensitivity of the results to changes in the parameter values, I conducted a one-way sensitivity analysis using the SCDD-I model. The value of each parameter was varied by 20%. In each of these analyses, the optimal solution remained the same as in the baseline analysis. The variation in expected profit is available in the supplementary material. For both vinblastine and vincristine, the unit and fixed cost values are most influential on expected profit. In particular, the unit cost of production has the largest effect on expected profit. The disruption and recovery distribution parameters have less of an impact.

I also conducted scenario analyses on the other parameters. These included the lengths of the time horizon and periods, annual demand, and production capacity. The optimal solution did not change for either drug though there were minor differences in expected profit. Further detail is available in the supplementary material.

To validate the models, I compared the results with available data in the literature. The

solutions at baseline also follow the lean supply chains and low inventory seen in practice (Fox et al. 2014, GAO 2016, Woodcock and Wosinska 2013). The failure-to-supply results are qualitatively confirmed with Jia and Zhao (2017). The baseline shortages of 6% for vinblastine and 11% for vincristine are similar to the percentage of drugs short each day reported by the drug shortage staff at a large academic health system. The other results have face validity; as price increases, resiliency increases, and the results mimic the dynamics of higher-margin, branded drugs.

7. Discussion

7.1. Modeling

Drug shortages are concerning because they are widespread, harmful, and persistent. To study why pharmaceutical companies may make supply chain decisions that contribute to shortages, I develop two new supply chain design models: SCDD and SCDD-I. These models combine features previously considered separately to provide a framework for understanding the effects of disruptions over time and for evaluating policies. They incorporate the multi-period aspect of inventory models under disruption with the facility selection decisions of location models to consider multiple mitigation strategies. In addition, disruptions may occur at multiple echelons and concurrently. These models allow us to approximate the strategic decisions pharmaceutical companies make for fixed-term contracts.

The baseline model, SCDD, is relatively simple and could be easily extended to include additional echelons, location decisions, or correlations between component disruptions. It is appropriate for settings in which inventory either cannot be held or is very expensive. The extended model, SCDD-I adds inventory as a resiliency strategy. This feature complicates the model, though I impose a replenishment rule that implies the non-anticipativity property holds in

the optimal solution. This substantially reduces the computational burden and allows us to use SAA to solve thirteen-stage stochastic programs within tight optimality gaps (i.e., thirteen based on the initial stage and 12 subsequent periods). Without this rule, the problem would require specialized algorithms to solve.

7.2. Is low resiliency optimal?

Using these models, I consider the case examples of the supply chains of the oncology drugs vinblastine and vincristine. While resiliency is often optimal in other contexts (Tomlin 2006), pharmaceutical companies may find instead that passive acceptance of risk is optimal for certain drugs, i.e., low-margin products with long, infrequent disruptions. The results suggest that with a profit-maximizing objective and no intervention, it would be best to have no resiliency in the supply chain of vincristine, a low volume drug, leading to an expected shortage of 11%. For the higher volume drug, vinblastine, it is beneficial to have a backup supplier, and the expected shortage is 6%. For other stakeholders in the healthcare system, these levels are untenable.

7.3. How can we induce resiliency and reduce drug shortages?

Given that it is in society's best interest to reduce shortages of life-supporting drugs, the question becomes which strategies to induce resiliency would be best. I use the metric of social-efficiency (i.e., lowest total cost to meet specified expected shortage levels) to evaluate proposed options. I studied the legislative policies of mandating redundancy and safety stock and contractual policies of failure-to-supply penalties and price increases. For each, I evaluated the prices needed to maintain company profits at baseline levels. As a reminder, these results are presented within the context of low-profit margin, sole-source, generic, injectable oncology drugs.

I find that the most efficient policy depends on the desired shortage level. For shortages of 5% or less (i.e., expected to be available at least 95% of the time), no intervention is needed for vinblastine, and for vincristine, it is most efficient to require multiple suppliers and increase prices by 10%. A failure-to-supply penalty of 70% would equivalently induce a backup supplier but would require a greater price increase, 20%. These results suggest that maintaining multiple suppliers is an effective way to reduce shortages, though the societal cost to induce them would depend on the decision maker.

Shortages of at most 2% could be achieved by requiring a backup plant with a 10% price increase (vinblastine) or requiring redundancy at all levels with a 30% price increase (vincristine). In both cases, the outcome is a supply chain with a backup at each echelon. An alternative policy to induce the same supply chain for vinblastine would be a failure-to-supply penalty of 150% with a price increase of 10%. For vincristine, an alternative would be to mandate the company hold one year of safety stock in combination with a 30% price increase.

In general, requiring safety stock is a relatively expensive policy option. This may be because the average disruption length is long and holding inventory of injectable drugs is costly. For vinblastine to have at most 2% shortages, a safety stock mandate is three times as costly as mandating multiple plants or adding a 150% failure-to-supply penalty (price increases of 30% vs. 10%). For vincristine, it is two times as costly to mandate sufficient inventory as it is to require multiple suppliers for a 5% shortage level (price increases of 20% vs. 10%). If safety stock were held, it could either be maintained at the manufacturer or in a stockpile similar to the Strategic National Stockpile of pediatric vaccines (Jacobson et al. 2006); the analysis would be the same in either case.

While shortages may be driven by low profit margins, in no analysis are price increases

alone the most efficient policy. For vinblastine, at the 2% shortage level, price increases are 15 times more costly than the most efficient policy (150% increase vs. 10%). For vincristine at the 5% shortage level, price increases are 7 times more costly (70% increase vs. 10%) than the most efficient option, and they are 27 times more costly (800% increase vs. 30%) at the 2% shortage level. These pricing results are consistent with an analysis from Jia and Zhao (2017) that found that adding failure-to-supply clauses in combination with moderate price increases would be more efficient than price increases alone.

These results suggest that legislative action to mandate redundant components in combination with price increases would have the lowest societal cost for both drugs. For vinblastine, sufficient failure-to-supply penalties in combination with price increases would also have the lowest societal cost; for vincristine, this is a costlier option. Though legislative change is difficult, it may be possible. A 2012 law changed reporting requirements for shortages, and in 2018, the FDA and other agencies initiated a Drug Shortage Task Force to provide new recommendations to Congress (FDASIA, Public Law 112–144 2012, Gottlieb 2018a). Contractual changes could be negotiated by Group Purchasing Organizations or other procurement officials. Price increases would likely be incurred by Medicare for Part B recipients and passed on to private payers for patients with private insurance.

7.4. Limitations

The results of this chapter are tempered by its limitations, and readers should be careful to interpret analyses within the appropriate scope. The models are subject to a variety of assumptions, discussed in Sections 3.4 and 4.4. The analyses assume a stationary market share, i.e., other companies do not enter the market. Given the high utilization of existing manufacturing capacity, most firms make decisions for a portfolio of products, rather than

individual drugs. These analyses assume the manufacturer does not choose to use the capacity for a more profitable drug. Finally, the analyses are limited by available data as pharmaceutical data are frequently proprietary. I have taken strides to estimate reasonable parameter values and conduct sensitivity analyses. In particular, profit results should not be taken as exact projections but rather as indications of the magnitude of policy effects. For this reason, the focus has been on the change vs. baseline rather than absolute numbers. The optimal supply chain configurations and target inventory levels do not vary as parameter values are varied by 20%, and analyses of demand indicate that the solutions do not change within wide ranges.

7.5. Conclusions

Strategic supply chain decisions have contributed to major drug shortages, and I find that for certain types of drugs with low profit margins, pharmaceutical companies may find it optimal to maintain vulnerable supply chains. In this analysis, I seek to align the interests of for-profit companies with the public good of a stable drug supply. Experts have suggested that regulation may be required to reduce shortages. The results provide evidence that redundancy regulations would be at least as efficient as market-based solutions. If legislation is pursued, additional analysis would be necessary to determine the particular characteristics of medically-necessary drugs to which it should be applied. In the absence of expanded regulation, group purchasing organizations and other contract-makers could negotiate failure-to-supply clauses in combination with modest price increases to reduce shortages. Price increases alone could also be effective but would cost substantially more. These models provide a framework to consider disruptions in strategic design decisions. Future work could consider improving quality as a resiliency strategy, which could lead to less frequent disruptions or faster recovery.

CHAPTER III

Dynamic Supply Chain Design under Disruption: Applications in Pharmaceutical Drug Shortages

1. Introduction

Drug shortages have become common in the United States (US), and they are largely caused by disruptions to non-resilient pharmaceutical supply chains (GAO 2014, UUDIS 2016). Facilities and suppliers may be temporarily closed due to quality concerns. Manufacturing delays, caused by capacity constraints and the high utilization of the manufacturing lines, may also interrupt production. Optimized supply chains may have little redundancy, and supply shocks can cause widespread shortages (Woodcock and Wosinska 2013).

Shortages persist for over a year on average, (GAO 2016) and often involve drugs produced by a single company (UUDIS 2016). Low profit margin, generic injectable drugs are particularly vulnerable to shortage. Nearly all of the manufacturers did not have an operational back-up facility if there was a disruption (Woodcock and Wosinska 2013).

If a disruption occurs, a pharmaceutical company could work to recover the disrupted component and/or add a new one. Yet, supply chains are fairly rigid, and it can take months for reviews of new components to be completed and to become available (GAO 2016). Because of this, previous models considered static designs to evaluate policies to reduce shortages (Chapter 2).

However, if the lead time to add new components or the recovery time of disrupted components were decreased, it could make the supply chain more flexible to adapt to disruptions. In this chapter, I present a dynamic version of the supply chain design model under disruption (D-SCDD). It will consider how the times to start up new components and recover disrupted components affect shortages. In addition, I will consider discontinuations due to disruption and consider how these parameters affect a company's decision to remain in the market. I will use a multi-stage stochastic programming framework and apply the Stochastic Dual Dynamic Integer Programming algorithm to solve the dynamic problem (Zou et al. 2019).

2. Literature

2.1. Drug shortages

While there have been numerous reports documenting the extent of the drug shortage crisis, there have been few quantitatively-based recommendations. In a working paper, Kim and Scott Morton develop a game-theory model to determine the extra manufacturing capacity that pharmaceutical companies should maintain for generic sterile injectable drugs (Kim and Scott Morton 2015). They find the Nash equilibrium of the capacity game where firms set spare capacity to maximize the payoff and provide the threshold for which firms should have a non-zero amount of spare capacity. There has been work to analyze policies to reduce shortages using Pareto-improving contracts (Jia and Zhao 2017) and optimizing static supply chain design (Chapter 2). Other work includes the study of human behavior (Doroudi et al., 2018) and reliability indices (Chapter 4). Much of the modeling work has focused on the prevention phase rather than reactive decisions. While there are mitigation strategies at the health system-level (Azghandi et al. 2018, Saedi et al. 2016), the focus of this chapter is to improve structural issues

that contribute to companies' inability to reliably provide supply. There is a need to study the dynamic context.

2.2. Disruptions

Beyond drug shortages, the operations research literature related to supply chains under disruption is rich and growing. A seminal paper by Kleindorfer and Saad (2005) divides disruption prevention into two categories: companies may either a) reduce the severity or incidence of risk, or b) add capacity to be able to sustain more risk. They propose that extreme leanness leads to vulnerability which is consistent with my observations in the pharmaceutical industry (Kleindorfer and Saad 2005). Within the context of infrastructure failures, both robustness (i.e., the ability to maintain full capacity) and the rapidity of recovery are important to reduce losses (McDaniels et al. 2008). The objective is not to maximize the former nor minimize the latter but rather to minimize the cumulative loss (to which both contribute).

A recent review of operations research papers related to disruptions provides references to a variety of modeling approaches and application areas (Snyder et al. 2016). Most supply chain design models that incorporate disruptions consider a single stage. These include facility location models under disruption (Snyder et al. 2006). A popular approach is to consider an interdiction framework where the objective is to minimize the maximum damage caused by an attacker (e.g., Church, Scaparra, and Middleton, 2004; O'Hanley and Church, 2011). This framework allows the company to prepare for the worst-case disruption that could occur. However, the perspective of interdiction models is inherently risk-averse, and the results would be very conservative for risk-neutral pharmaceutical companies.

Among multi-period disruption models, a simulation evaluated the effects of disruption risk in a high-tech supply chain over the course of three months (Deleris and Erhun 2005). A

risk-averse facility location model considered the question of recovery time (Losada et al. 2012), and a multi-stage stochastic program considered demand that was dependent on facility availability (Fattahi et al. 2017).

A recent review presented an overview of supply chain network design problems under uncertainty (Govindan et al. 2017). Only a handful of papers have considered disruptions and included more than one echelon. These have included a robust optimization approach to handling unreliable nodes (Peng et al. 2011) and constraining the maximum allowable disruption cost (Shishebori et al. 2014). An early work minimized the expected costs of constructing the network such that service level constraints were satisfied (Bundschuh et al. 2003). Hopp and Yin (2006) considered where to install extra capacity in a supply chain network subject to disruptions.

Modelers also generally assume that disruptions are independent across periods. That is, they may consider yield to be a Bernoulli random variable (Snyder et al. 2016) which allows for stage-wise independence (e.g., Fattahi, Govindan, and Keyvanshokooh, 2017).

However, in the context of drug shortages, recovery is long, and the approval process of new suppliers and plants can extend far beyond the length of single periods. Research on re-configuring supply chains with uncertainty over time is rare (Govindan et al. 2017), and this chapter helps to address this gap. I consider both the dynamic case and dependent disruptions.

2.3. Capacity adjustment

When conditions change over time – be it demand, capacity availability, or another parameter – organizations may adjust their strategic decisions. Within the facility location literature, the key choices are when and where to add capacity and whether capacity should be removed. Reviews of the capacity expansion literature are available in (Julka et al. 2007, Luss

1982, Martínez-Costa et al. 2014). Most commonly, capacity is added to meet demand. These include a multi-period SCND model with demand and interest rate uncertainty (Nickel et al. 2012) and a multi-product facility location model with increasing, deterministic demands (Thanh et al. 2008).

When capacity is added, there may be a lead time between the decision to add it and its ability to be used. This has been included in some models, especially when the model considers multiple sites (e.g., Lin et al. 2014).

2.4. Multi-stage stochastic programming

Multi-stage stochastic programs have been used in some SCND problems that include uncertainty (Govindan et al. 2017), though the problems studied have tended to be small (e.g., 9 scenarios Almansoori and Shah 2012). Multi-stage stochastic programs have been used broadly in other contexts, however, and a textbook treatment is available in Birge and Louveaux (2011).

Large-scale linear multi-stage stochastic programs are often solved using variants of the Stochastic Dual Dynamic Programming (SDDP) algorithm (Pereira and Pinto 1991). The algorithm considers a nested multi-stage stochastic program; the objective function of each stage is based on the costs incurred at the current stage and an approximation term of the expected value-to-go. A dynamic programming recursion iteratively passes the scenario tree forwards and backwards to generate cuts for the approximation of the expected value-to-go. Though widely used, the algorithm requires uncertainty to be stage-wise independent and decision variables to be continuous. This would not apply to the model proposed in this chapter; the uncertainty I consider is stage-wise dependent with binary decision variables.

Recent work has extended SDDP to address these restrictions. Philpott and de Matos (2012) consider cases where the uncertainty is defined according to a hidden, discrete-time

Markov chain. They test instances with four states in each stage but suggest that their algorithm would not scale well beyond 16 states. Given a set of N candidate components, my model would have 2^|N| states in each stage, i.e., easily over 16. A more promising algorithm for this context is the Stochastic Dual Dynamic Integer Programming (SDDiP) algorithm (Zou et al. 2019). This method extends SDDP to allow decision variables to be binary. Yet, it requires stage-wise independent uncertainty. This is because the cuts are shared across nodes in a stage rather than maintained at each node.

As the data on the times to disruption and recovery in the pharmaceutical industry follow a geometric process (FDA 2018a, UUDIS 2016), sampling the component statuses would typically be modeled using stage-wise dependence. Researchers have developed cut-sharing methods to adjust cuts to be used at multiple nodes in the same stage (Infanger and Morton 1996, De Queiroz and Morton 2013). In other cases, modelers can transform the underlying stage-wise dependent stochastic process to be an autoregressive process with an independent error term and apply the methods as presented (Shapiro 2011).

The geometrically-distributed times to recover and disruption do not have random errors. Instead, I reformulate the Markovian random variables to be stage-wise independent by sampling realizations of the inverse cumulative distribution function (CDF) of the appropriate geometric distributions. Then, I apply the SDDiP algorithm to the dynamic supply chain design problem.

2.5. Contributions

The contributions of this chapter are the following:

- I present a dynamic design model where the supply chain may be disrupted and recover over time. The company may choose to re-configure its supply chain over time.

- The model is used to study the effects of disruptions on configuration changes and production discontinuations.
- I apply the SDDiP algorithm to supply chain design under disruption and reformulate the Markovian random process to be stage-wise independent.
- The analyses present the effects of varying the lead time of starting up a new component and improving the recovery process.

3. Model

3.1. Overview

The model determines the dynamic supply chain configuration when each component is at risk of disruption. Supply chain components include raw material suppliers, manufacturing plants, and manufacturing lines within the plants. Each may be available or disrupted in a given period. Demand is deterministic and constant across the time horizon.

The model represents a single firm making choices about a single drug. The objective is to maximize the expected profit under uncertainty regarding the status of the supply chain components. The model is formulated as a multi-stage stochastic program. In each stage, the company decides which candidate supply chain components to use, how much to order from the raw material suppliers, production quantities at the manufacturing plants, and how much to sell (Figure 9). At the initial selection ($t = 1$), the company does not know which components will be working in the subsequent stages. In the second stage ($t = 2$), the component statuses are sampled from the steady-state distribution of availability. Uncertainty is realized progressively at each stage, and demand is either met or there is a shortage.

The supply chain configuration may be changed over the course of the time horizon. If the company is in the market, the company must maintain the components that are selected in the

initial stage. It may start the process of adding new components at any stage, given eligibility criteria are met. Components can only be added when there is a disrupted component in the echelon. As a simplifying assumption, new lines may be added within new plants that are added to the configuration but not within plants that are already part of the configuration. The latter restriction is due to the quick recovery of the lines.

There is a lead time to add new components. For example, a one stage lead time would require the company to wait one stage until the component may be used (Figure 10). Components that are added after the initial stage may be removed at any stage.

The company may also choose to discontinue the drug for all stages prior to a cutoff at which it may no longer leave the market. If it discontinues, the components are not available for the remainder of the horizon including the current stage; the company incurs no costs and generates no revenue.

The fixed costs include the annual fixed cost of components, the required annual Generic Drug User Fee Amendment (GDUA) fees for API suppliers and manufacturing plants, and the fixed costs during the lead time process. Variable costs are incurred due to production and raw material orders. The revenues come from drug sales.

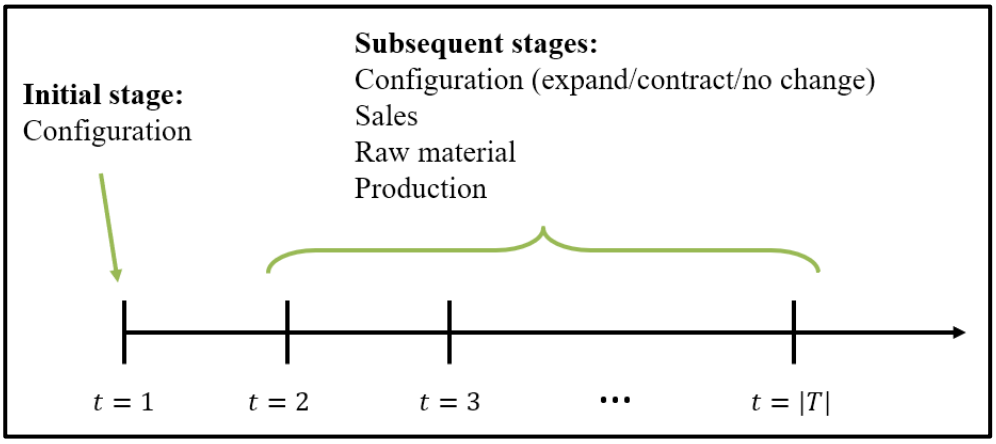

Figure 9. Timing of decisions

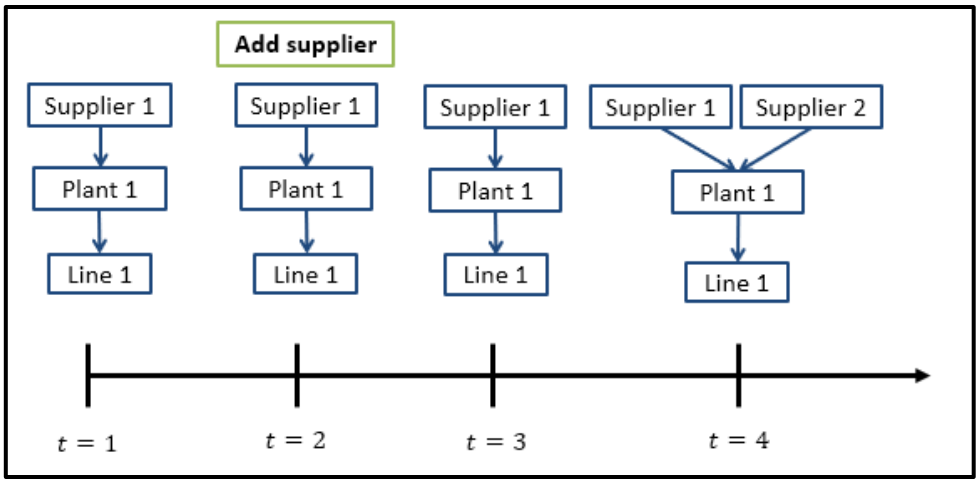

Figure 10. Expansion example

3.2. D-SCDD

In this section, I present the Dynamic Supply Chain Design Model under Disruption (D-SCDD). The sets and parameters are presented in Table 5. The decision variables are presented in Table 6. There are two primary types of variables: state and local. The state variables are those that are carried over to the next stage; the local variables are not recorded across stages. They are only used in the current stage.

Each state variable has a corresponding auxiliary variable, designated with a tilde. The auxiliary variables are set equal to the (fixed) state variables from the previous stage $t-1 \in T$. For example, within stage t, the state variable $x_{j,t-1}$ is fixed and is set equal to the auxiliary variable $\tilde{x}_{jt}$. These variables are redundant, but the dual variables of these constraints are used to generate cuts in the backwards recursion. This will be discussed in more detail in Section 4. For notational convenience, the variable z_t will be used to represent a vector of all of the state variables in $t \in T$.

D-SCDD is formulated as a nested stochastic program. The function $\psi_t(z_t)$ calculates the expected value-to-go. The objective of the sub-problem, $P_t(\psi_t(z_t))$, at each stage $t \in T$ is

to maximize the profit at the current stage plus an approximation of the expected value-to-go in future stages. The constraints are applied to the current stage t. Cuts are iteratively generated via the SDDiP algorithm to improve the approximation of the expected value-to-go (Section 4). Below, the formulation is presented by stage, and an explanation of each expression follows.

Table 5. Sets and parameters

Sets	
H	Set of echelons, i.e., $H \coloneqq \{API, Plant, Line\}$
J^h	Set of candidate components in echelon $h \in H$
J	Set of all components, where $J \coloneqq \cup_{h \in H} J^h$
$\ddot{t}$	Number of periods per year
T	Set of time periods, defined such that $\{\|T\| \ mod \ \ddot{t} = 1\}$, i.e., the length of the time horizon can be divided into years plus a initial period
I	Set of iterations
R_t	Set of realizations at stage $t \in T$
Parameters	
L^h	Number of periods required to open a new component in echelon $h \in H$
b_{lp}	$\coloneqq \begin{cases} 1 & \text{if candidate line } l \in J^{Line} \text{ is in plant } p \in J^{Plant} \\ 0 & \text{otherwise} \end{cases}$
d	Quantity of drug demanded per period
c^{raw}	Unit cost of raw material
c^{prod}	Unit production cost
f_h	Fixed costs per period per component in echelon $h \in H$
$f^{Program}$	Per period GDUFA program fee
c_h^{start}	Start-up costs per component in echelon $h \in H$
c_h^{LT}	Fixed costs per period per component of lead time to acquire a component in echelon $h \in H$
$f^{Program}$	Per period GDUFA program fee
q	Sales price per unit of drug
t^{nodisc}	Period at which company can no longer discontinue production and leave the market
p_j^{Fail}	Per period probability of component $j \in J$ failing
p_j^{Rec}	Per period probability of component $j \in J$ recovering
$p_j^{W,steady}$	Probability of component $j \in J$ working in steady-state
$\bar{\xi}_{j0}$	$\coloneqq \begin{cases} 1 & \text{if component } j \in J \text{ is initially working} \\ 0 & \text{otherwise} \end{cases}$
U_{jt}	Probability value corresponding to component $j \in J$ in stage $t \in T$
G_t	Upper bound on the objective function at stage $t \in T$
v_r^i	Cut coefficient for realization $r \in R_t$ for iteration $i \in I$

Table 6. Decision variables

State	Auxiliary	Description
x_{jt}	$\tilde{x}_{jt}$	$:= \begin{cases} 1 & \text{if component } j \in J \text{ is in the configuration in stage } t \in T \\ 0 & \text{otherwise} \end{cases}$
x_{jt}^{init}	$\tilde{x}_{jt}^{init}$	$:= \begin{cases} 1 & \text{if component } j \in J \text{ is open at the initial stage} \\ 0 & \text{otherwise} \end{cases}$
y_{jkt}	$\tilde{y}_{jkt}$	$:= \begin{cases} 1 & \text{if component } j \in J^h \text{ in echelon } h \in H \text{ has } k \in \{1, \dots, L^h\} \text{ periods} \\ & \text{remaining as of } t \in T \text{ until it can be added to the configuration} \\ 0 & \text{otherwise} \end{cases}$
ϕ_t	$\tilde{\phi}_t$	$:= \begin{cases} 1 & \text{if the company is in the market in stage } t \in T \\ 0 & \text{otherwise} \end{cases}$
ξ_{jt}	$\tilde{\xi}_{jt}$	$:= \begin{cases} 1 & \text{if component } j \in J \text{ is non} - \text{disrupted in stage } t \in T \\ 0 & \text{otherwise} \end{cases}$
z_t	$\tilde{z}_t$	The vector of all binary state variables in stage $t \in T$
Local		
δ_{jt}		$:= \begin{cases} 1 & \text{if component } j \in J \text{ maintains the same status from } t-1 \\ & \text{to } t \in T \text{ initeration } i \in I \\ 0 & \text{otherwise} \end{cases}$
τ_{jt}		$:= \begin{cases} 1 & \text{if component } j \in J \text{ is non} - \text{disrupted in } t-1 \in T \text{ and} \\ & \text{continues to be non} - \text{disrupted in } t \in \\ 0 & \text{otherwise} \end{cases}$
η_{jt}		$:= \begin{cases} 1 & \text{if component } j \in J^h \text{ is in the configuration yet disrupted} \\ & \text{in } t \in T \\ 0 & \text{otherwise} \end{cases}$
γ_t		Approximation of the expected value-to-go at stage $t \in T$
θ_t		Fraction of per period demand met in stage $t \in T$
u_{jt}		Raw material that is supplied by supplier $j \in J^{API}$ in stage $t \in T$ (as a fraction of per period demand)
w_{jt}		Finished goods produced on line $j \in J^{Line}$ in stage $t \in T$ (as a fraction of per period demand)
π_r^i		Cut coefficient for realization $r \in R_t$ in iteration $i \in I$ in stage $t \in T$

For $t = 1$:

$$\left(P_t(\psi_t)\right): \max \sum_{h \in H} \sum_{j \in J^h} -c_h^{start} x_{jt} + \psi_t(z_t) \tag{1}$$

Subject to:

$$x_{jt} \geq x_{j+1,t} \qquad \forall h \in H, j \in J^h \setminus \{|J^h|\} \tag{2}$$

$$x_{lt} + b_{lp} \leq 1 + x_{pt} \qquad \forall l \in J^{Line}, p \in J^{Plant} \tag{3}$$

$$x_{pt} \leq \sum_{l \in J^{line}} b_{lp} x_{lt} \qquad \forall p \in J^{Plant} \qquad (4)$$
$$\phi_t \leq \sum_{j \in J^h} x_{jt} \qquad \forall h \in H \qquad (5)$$

$$\phi_t \geq x_{jt} \qquad \forall j \in J \qquad (6)$$

$$x_{jt}^{init} = x_{jt} \qquad \forall j \in J \qquad (7)$$

$$\xi_{jt} = \bar{\xi}_{j0} \qquad (8a)$$
$$\theta_t = 1 \qquad (8b)$$
$$y_{jkt} = 0 \qquad \forall h \in H, j \in J^h, k \in \{1, \dots, L^h\} \qquad (8c)$$

$$x_{jt}, x_{jt}^{init}, \xi_{jt} \in \{0,1\} \qquad \forall j \in J \qquad (9a)$$
$$y_{jkt} \in \{0,1\} \qquad \forall h \in H, j \in J^h, k \in \{1, \dots, L^h\} \qquad (9b)$$
$$\phi_t \in \{0,1\} \qquad (9c)$$
$$\theta_t \geq 0 \qquad (9d)$$

Where:

$$\psi_t(z_t) \coloneqq \max \lambda_t \qquad (10)$$

Subject to:

$$\lambda_t \leq G_t \qquad (11)$$
$$\lambda_t \leq \frac{1}{|R_{t+1}|} \sum_{r \in R_{t+1}} \left(v_r^{i'} + \left(\pi_r^{i'}\right)^T z_t \right) \qquad \forall i' \in \{1, \dots, i-1\} \qquad (12)$$

For $\boldsymbol{t \in T \setminus \{1, |T|\}}$**:**

$\left(P_t(z_{t-1}, \psi_t)\right)$:

$$\max d\left(q\theta_t - c^{raw} \sum_{j \in J^{API}} u_{jt} - c^{prod} \sum_{j \in J^{Line}} w_{jt}\right) - \sum_{h \in H} \sum_{j \in J^h} \left[f_h x_{jt} + c_h^{LT} \sum_{k \in \{1, \dots, L^h\}} y_{jkt} \right] - f^{Program} \phi_t + \psi_t(z_t) \qquad (13)$$

Subject to:

$$x_{jt}^{init} = \tilde{x}_{jt}^{init} \qquad \forall j \in J \qquad (14)$$
$$x_{jt} \geq x_{jt}^{init} - (1 - \phi_t) \qquad \forall j \in J \qquad (15)$$

$$y_{j,k-1,t} = \tilde{y}_{jkt} \qquad \forall h \in H, j \in J^h, k \in \{2, \dots, L_h\} \qquad (16)$$
$$x_{jt} \leq \tilde{x}_{jt} + \tilde{y}_{j,1,t} \qquad \forall h \in H, j \in J^h \qquad (17)$$
$$\sum_{k \in \{1, \dots, L_h\}} y_{jkt} \leq 1 - \tilde{y}_{j,1,t} \qquad \forall h \in H, j \in J^h \qquad (18)$$

$$\eta_{jt} \leq x_{jt} \qquad \forall j \in J \qquad (19)$$
$$\eta_{jt} \leq 1 - \xi_{jt} \qquad \forall j \in J \qquad (20)$$

$$\eta_{jt} \geq x_{jt} + \left(1 - \xi_{jt}\right) - 1 \qquad \forall j \in J \qquad (21)$$

$$y_{j,L^h,t} \leq \sum_{j \in J^h} \eta_{jt} \qquad \forall h \in H, j \in J^h \qquad (22)$$

$$x_{lt} + b_{lp} \leq 1 + x_{pt} \qquad \forall l \in J^{Line}, p \in J^{Plant} \qquad (23)$$

$$x_{pt} \leq \sum_{l \in J^{line}} b_{lp} x_{lt} \qquad \forall p \in J^{Plant} \qquad (24)$$

$$y_{lkt} + b_{lp} \leq 1 + y_{pkt} \qquad \forall p \in J^{Plant}, l \in J^{Line}, k \in \{2, \ldots, L^{Plant}\} \qquad (25)$$

$$y_{pkt} \leq \sum_{l \in J^{Line}} y_{lkt} b_{lp} \qquad \forall p \in J^{Plant}, k \in \{2, \ldots, L^{Plant}\} \qquad (26)$$

$$\phi_t \leq \sum_{j \in J^h} x_{jt} \qquad \forall h \in H \qquad (27)$$

$$\phi_t \leq \tilde{\phi}_t \qquad (28)$$

$$\phi_t \geq x_{jt} \qquad \forall j \in J \qquad (29)$$

$For: t \geq t^{nodisc}$

$$\phi_t \geq \tilde{\phi}_t \qquad (30)$$

$$u_{jt} \leq x_{jt} \qquad \forall j \in J^{API} \qquad (31)$$

$$u_{jt} \leq \xi_{jt} \qquad \forall j \in J^{API} \qquad (32)$$

$$\sum_{l \in J^{Line}} w_{lt} \leq \sum_{j \in J^{API}} u_{jt} \qquad (33)$$

$$w_{lt} \leq \sum_{p \in J^{Plant}} b_{lp} \xi_{pt} \qquad \forall l \in J^{Line} \qquad (34)$$

$$w_{jt} \leq \xi_{jt} \qquad \forall j \in J^{Line} \qquad (35)$$

$$w_{jt} \leq x_{jt} \qquad \forall\, j \in J^{Line} \qquad (36)$$

$$\theta_t \leq 1 \qquad (37)$$

$$\theta_t \leq \sum_{j \in J} w_{jt} \qquad (38)$$

$t = 2$:

$$\left(1 - p_j^{W,steady}\right) - \left(1 - U_{jt}\right) \leq 1 - \xi_{jt} \qquad \forall j \in J \qquad (39)$$

$$p_j^{W,steady} - U_{jt} \leq \xi_{jt} \qquad \forall j \in J \qquad (40)$$

Otherwise:

$$\xi_{jt} = \tilde{\xi}_{jt} \delta_{jt} + \left(1 - \tilde{\xi}_{jt}\right)\left(1 - \delta_{jt}\right) \qquad \forall j \in J \qquad (41)$$

$$\tau_{jt} \leq \tilde{\xi}_{jt} \qquad \forall j \in J \qquad (41a)$$

$$\tau_{jt} \leq \delta_{jt} \qquad \forall j \in J \qquad (41b)$$

$$\tau_{jt} \geq \tilde{\xi}_{jt} + \delta_{jt} - 1 \qquad \forall j \in J \qquad (41c)$$

$$\xi_{jt} = 1 - \tilde{\xi}_{jt} - \delta_{jt} + 2\tau_{jt} \qquad \forall j \in J \qquad (41d)$$

$$\left[p_j^{Fail}\tilde{\xi}_{jt} + p_j^{Recov}\left(1-\tilde{\xi}_{jt}\right)\right] - \left(1-\delta_{jt}\right) \leq U_{jt} \qquad \forall j \in J \qquad (42)$$

$$\left[p_j^{Fail}\tilde{\xi}_{jt} + p_j^{Recov}\left(1-\tilde{\xi}_{jt}\right)\right] + \delta_{jt} \geq U_{jt} \qquad \forall j \in J \qquad (43)$$

$$\tilde{x}_{jt} = x_{j,t-1} \qquad \forall j \in J \qquad (44a)$$

$$\tilde{x}_{jt}^{init} = x_{j,t-1}^{init} \qquad \forall j \in J \qquad (44b)$$

$$\tilde{y}_{jkt} = y_{j,k,t-1} \qquad \forall j \in J^h, h \in H, k \in \{1,\dots,L^h\} \qquad (44c)$$

$$\tilde{\phi}_t = \phi_{j,t-1} \qquad (44d)$$

$$\tilde{\xi}_{jt} = \xi_{j,t-1} \qquad \forall j \in J \qquad (44e)$$

$$x_{jt}, \tilde{x}_{jt}, x_{jt}^{init}, \tilde{x}_{jt}^{init}, \xi_{jt}, \tilde{\xi}_{jt}, \delta_{jt}, \tau_{jt}, \eta_{jt} \in \{0,1\} \qquad \forall j \in J \qquad (45a)$$

$$y_{jkt}, \tilde{y}_{jkt} \in \{0,1\} \qquad \forall j \in J^h, h \in H, k \in \{1,\dots,L^h\} \qquad (45b)$$

$$\phi_t, \tilde{\phi}_t \in \{0,1\} \qquad (45c)$$

$$u_{jt} \geq 0 \qquad \forall j \in J^{API} \qquad (45d)$$

$$w_{jt} \geq 0 \qquad \forall j \in J^{Line} \qquad (45e)$$

$$\theta_t \geq 0 \qquad (45f)$$

$$\beta_{ht} \in \{0,1\} \qquad \forall h \in H \qquad (45g)$$

$$U_{jt} \geq 0 \qquad \forall j \in J \qquad (45h)$$

Where:

$$\psi_t(z_t) \coloneqq \max \lambda_t \qquad (46)$$

Subject to:

$$\lambda_t \leq G_t \qquad (47)$$

$$\lambda_t \leq \frac{1}{|R_{t+1}|}\sum_{r\in R_{t+1}}\left(v_r^{i'} + \left(\pi_r^{i'}\right)^T z_t\right) \qquad \forall i' \in \{1,\dots,i-1\} \qquad (48)$$

Objective $(\boldsymbol{t = |T|})$**:**

$\left(P_t(z_{t-1})\right)$:

$$\max d\left(q\theta_t - c^{raw}\sum_{j\in J^{API}} u_{jt} - c^{prod}\sum_{j\in J^{Line}} w_{jt}\right) - \sum_{h\in H}\sum_{j\in J^h}\left[f_h x_{jt} + c_h^{LT}\sum_{l\in\{1,\dots,L^h\}} y_{jlt}\right] - f^{Program}\phi_t \qquad (49)$$

Subject to:

(14-45)

Expressions (1-12) define the initial stage problem. Expressions (13-49) define the problem for the subsequent stages.

In the initial stage, $t = 1$, the objective function (1) maximizes expected profit by subtracting the fixed start-up costs for each component and adding an approximation of the expected value-to-go. The approximation is defined via the sub-problem in expressions (10-12).

Constraints (2) require the components to be selected in the order of their indices. These are used to reduce the number of alternative optima. Constraints (3) require that the plant associated with a selected line is selected. Constraints (4) do not allow the company to select a manufacturing plant unless it contains a selected line. Constraints (5-6) determine whether the company is in the market. Constraints (5) require a component to be selected in each echelon. Constraints (6) ensure that if a component is selected, the company is considered in the market. The combination of (5-6) requires that there is either a complete supply chain configuration or no components selected. Constraints (7) record the initial supply chain configuration.

Constraints (8) initialize the component statuses, demand met, and lead time variables, respectively. Constraints (9) enforce variable domains.

Expressions (10-12) are the sub-problem that defines the approximation of the expected value-to-go. Objective (10) maximizes the value-to-go approximation. Constraint (11) limits the approximation to an exogenous upper bound. Constraints (12) are cutting planes that provide an outer-approximation of the value-to-go. The coefficients $\{v_r^{i'}, \pi_r^{i'}\}, \forall r \in R_{t+1}, i \in \{1, \dots, i-1\}$ are fixed in constraints (12) and are defined by solving the relaxation $\left(R_t^i\right)$. This is discussed in Section 4.

In the subsequent stages, where $t \in \{2, .., |T| - 1\}$, the objective function (13) maximizes the sum of the profit in the current stage and an approximation of the expected value-to-go. The profit is the sum of revenue generated from drug sales, the cost of ordering raw materials, the cost of producing finished goods, and the fixed costs. There are fixed costs for

each component and the GDUFA program fee. The approximate value-to-go is calculated via the sub-problem in expressions (46-48).

Constraints (14-26) focus on the selected configuration and lead time. Constraints (14) record the supply chain configuration selected in the initial stage. Constraints (15) require the initial configuration to be maintained unless the company is not in the market.

Constraints (16-18) enforce the lead time to add a new component. The lead-time-remaining indicator records how many stages remain before a component can be added to the configuration. Constraints (16) decrease the indicator by one period from the previous stage. Constraints (17) require that each component in the configuration either i) to have been in the configuration in the previous stage or ii) to have had one period of lead time remaining in the previous stage. Constraints (18) force the lead-time-remaining indicator to 0 if there was one period remaining in the previous period. It also limits each component to be at a single point in the lead time process at any given time. Multiple components can have different lead times, but each component can only have one lead time in a given period.

Constraints (19-22) require that a company may only begin the lead time of a component if a component within the echelon is in the configuration and non-disrupted. Constraints (19-21) define the component selected-and-available variable. Constraints (22) require that the process to add a new component in echelon $h \in H$ may only begin if there is a disrupted component in the echelon.

Constraints (23-26) enforce plant-line compatibility. Constraints (23) and (25) require lines must be in selected plants, for selected components and the lead-time process, respectively. Constraints (24) and (26) ensure at least one line is selected in a selected plant, for selected

components and the lead-time process, respectively. Constraints (25) ensure that when the process to add a new plant begins, at least one line within the new plant will also be added.

Constraints (27-30) relate to the company's status in the market. Constraints (27) require at least one component at each echelon to be selected if the company is in the market. Constraint (28) does not allow the company to reenter the market if they have exited. Constraints (29) designate the company to be in the market if any components are selected. Beginning at period t^{nodisc}, constraint (30) requires the company to be in the market if they were in the previous period.

Constraints (31) limit the raw materials purchased to selected API suppliers. Constraints (32) prevent raw materials from being purchased from disrupted API suppliers. Constraints (33-36) impose upper bounds on the amount of finished goods that can be produced. Constraints (33) limit the total amount of drugs produced to the amount of raw materials available. Constraints (34-35) prevent the drug from being produced in disrupted plants and on disrupted lines, respectively. Constraints (36) limit the amount of drug produced to selected lines.

Constraints (37) prevent the company from selling more drug than is demanded. Constraints (38) limit the amount of drug sold to the amount that has been produced in the period. The company may not carry any inventory.

In period $t = 2$, constraints (39-40) define the status of each component. They are based on the steady-state availability of each component.

For all other periods, Constraints (41) define the status of each component $j \in J$ in the current stage $t \in T \setminus \{1\}$ based on: i) the status in the previous stage, $\tilde{\xi}_{jt}$, and ii) the indicator variable, δ_{jt}, that designates whether it maintains the same status. These are non-linear

expressions because they include products of binary variables, and constraints (41a-d) are used to implement an extra linearization.

Constraints (42-43) define the indicator of whether a component maintains the same status from the previous stage. The expression $\left[p_j^{Fail}\tilde{\xi}_{jt} + p_j^{Recov}\left(1-\tilde{\xi}_{jt}\right)\right]$ represents the probability that a component changes its status. The binary variables δ_{jt} indicate whether the component $j \in J$ maintains the same status in stage $t \in T$. U_{jt} are realizations of the uniform random variable corresponding to component $j \in J$ in stage $t \in T$. They are sampled before the model is run. That is, $\delta_{jt} = 0$ if $U \leq \left[p_j^{Fail}\tilde{\xi}_{jt} + p_j^{Recov}\left(1-\tilde{\xi}_{jt}\right)\right]$, and 1 otherwise. The indicator is defined using Big-M style constraints (39-40) with $\mathrm{M} = 1$.

In more detail:

If the status of component $j \in J$ in the previous stage is non-disrupted ($\tilde{\xi}_{jt} = 1$), the constraints are simplified to:

$$p_j^{Fail} - \left(1 - \delta_{jt}\right) \leq U_{jt} \qquad \forall j \in J$$

$$p_j^{Fail} + \delta_{jt} \geq U_{jt} \qquad \forall j \in J$$

The probability the component is disrupted is p_j^{Fail}. If the realization of the uniform random, U_{jt}, is less than p_j^{Fail}, then the indicator variable δ_{jt} is forced to 0. If it is greater than p_j^{Fail}, then the indicator variable is forced to 1. This represents a probability of p_j^{Fail} that component $j \in J$ changes status if it is non-disrupted in the previous period. The analysis is similar if the component is disrupted in the previous stage, i.e., $\tilde{\xi}_{jt} = 0$.

Constraints (44a-e) assign the auxiliary state variables as the fixed values of the state variables from the previous stage. Constraints (45a-h) enforce the domains.

Expressions (46-48) calculate the approximation of the expected value-to-go. They are similar to expressions (10-12) in the initial stage. The objective (46) maximizes the approximation. Constraint (47) limits it to a given upper bound. Constraints (48) apply cuts based on fixed coefficients $\{v_r^{i'}, \pi_r^{i'}\}, \forall r \in R_{t+1}, i \in \{1, \dots, i-1\}$.

When $t = |T|$, the objective function (49) maximizes the profit in the final stage. It is calculated in the same way as expression (13) except it excludes an approximation of the expected value-to-go. The model is subject to the same constraints as the other non-initial stages (14-45).

Note that none of the stages include non-anticipativity constraints. These are typically applied in multi-stage stochastic programs that are decomposed by scenario (rather than by stage). They prevent the model from incorrectly applying information about future stages. D-SCDD is modeled using a nested structure, and the non-anticipativity property is implied.

3.3. Assumptions

The model includes several assumptions. The API suppliers represent individual facilities and can supply any plant. As freight shipping between common supplier and plant locations is sufficiently shorter than the length of a period (e.g., 2 weeks vs. three months), (SeaRates 2018, US Department of Commerce 2018), the shipping time between a supplier and plant is assumed to be zero. The transportation costs are not explicitly modeled, though they could be implicitly included via the raw material ordering costs. As a simplifying assumption, the GDUFA fees for API suppliers and facilities are applied per stage rather than annually. They are only applied to components that are in the configuration, not those in the expansion process. A separate per-period cost to the company is applied to the latter. I assume each non-disrupted

line can produce sufficient quantities to meet demand and each non-disrupted API supplier can supply sufficient raw materials.

It takes time to add a new component, and I assume that this lead time is deterministic. In practice, this would vary, but I applied a known, constant value to allow the analysis to focus on the effects of varying the lead time. In contrast, disruptions and recovery are uncertain. They occur to components independently, and components are either available or disrupted. Within an echelon, each component has the same probability of disruption recovery each period. The model does not consider correlations nor partial availability, following the analysis in earlier work (Chapter 2).

The demand is deterministic and constant, similar to the demand distributions of many drugs vulnerable to shortages (Fox et al. 2014). If demand is unmet, it is lost, not backordered. The sales price and costs are constant across time periods. Finally, the company is not allowed to hold inventory. Companies of generic sterile injectables typically hold low amounts of inventory (GAO 2016).

4. Solution methods

To solve D-SCDD, I apply the SDDiP algorithm, introduced by Zou et al. (2019). It is appropriate for multi-stage stochastic mixed-integer linear programming problems. The objective at each stage is the profit at the current stage and the expected value-to-go in the remaining stages. The algorithm decomposes the model by stage. The objective then becomes the profit at the current stage plus an approximation term of the expected value-to-go. To improve the approximation, each iteration of the algorithm generates cutting planes that form an outer-approximation of the value-to-go term. The sub-problems are much easier to solve than

the full problem, and the algorithm progressively traverse stages of the tree to generate cuts and approximate the optimal solution of the full problem.

Each iteration has two steps – a forward pass and a backwards pass. In the forward step, the algorithm samples one path of realizations of uncertainty, $U_{jt},\ \forall j \in J, t \in T$, corresponding to each component j and each time period t. Then from stage $t = 1$ to $t = |T|$, it solves the optimization problem $P_t^i(\psi_t^i)$ in each stage. It records the state variables, z_t, and value-to-go approximation, λ_t, for each stage.

In the backward step, the algorithm traverses the sample path "backwards" in time, from $t = |T|$ to $t = 2$. At each included stage, the algorithm solves a relaxation of the stage problem (discussed more specifically in Section 4.3), and the cut coefficients are obtained.

The algorithm iterates through forward and backward steps. It terminates when the objective function is within appropriate optimality bounds or a maximum number of iterations is reached.

4.1. Applicability of the algorithm

The SDDiP algorithm requires the following conditions. The problem must i) have a linear objective function in the state and local variables at each state; ii) have a nonempty, compact, mixed-integer polyhedral set; and iii) have complete, continuous recourse.

These conditions hold for D-SCDD. The objective function is linear in the variables at each stage. The constraint set is nonempty; it is always feasible to select no components. It is compact (closed and bounded) because a) none of the inequalities are strict, and b) each variable is limited below by 0 and has an upper bound. The binary variables have an upper bound of 1, and the continuous variables (i.e., raw materials ordered, production quantities) are limited by capacity. Capacity is bounded by the discrete, finite set of candidate components. Demand met

is limited by demand. Finally, it has complete, continuous recourse because for any value of the state variables (i.e., $x_{jt}, x_{jt}^{init}, y_{jkt}, \phi_t, \xi_{jt}$), the local continuous variables (i.e., u_{jt}, w_{jt}) can be set to 0.

4.2. Uncertainty

The SDDiP algorithm requires the uncertainty to be stage-wise independent. This is relevant to many supply chain models that incorporate disruptions. Often, the availability of components is modeled using a Bernoulli random variable that does not depend on prior periods (Snyder et al. 2016). However, if disruptions occur in the pharmaceutical industry, they may persist for many months, and the times to disruption and recovery follow different geometric distributions (Chapter 2).

When the times to disruption and recovery have different distributions, the component statuses, i.e., the realizations of uncertainty, are no longer stage-wise independent. As an example, if geometrically distributed disruptions are modeled using the conventional framework where availability is the random variable, the status at stage $t-1$ affects the status at stage t. That is, if a component is disruption in stage $t-1$, it is more likely to still be disrupted in stage t than if it were available in stage t.

In other contexts, a model can sometimes be reformulated to have stage-wise independent uncertainty. This is generally done by converting the random process to an autoregressive process (Shapiro 2011). The stage-wise independent random errors are sampled, and the random variables corresponding to the original process can be recovered using the constraints. Each node of the scenario tree is a realization of the random error, rather than a realization of the variables from the original process.

In a Markov process, there are not random errors. However, inspired by this reformulation idea, I redefine the random variables in two steps. First, I add binary variables to represent whether the component maintains its status from the previous stage (i.e., 1 if it maintains the same status from period $t-1$ to t, and 0 if it changes status). The second step is to add random variables that represent the realization of probability it changes its status i.e., $U_{jt} \sim Unif[0,1], \forall\, j \in J, t \in T$. These are uniform random variables distributed between 0 and 1. This process essentially samples the inverse of the appropriate cumulative distribution function (CDF) of the geometric distribution. Yet, it is stage-wise independent because the appropriate inverse CDF is applied in the constraints, rather than the uncertainty realizations. The random variable itself does not depend on the previous stage; it simply samples $U \sim Unif[0,1]$.

Constraints (41-43) use the realizations U_{jt} and the exogenous parameters p_j^{Recov} and p_j^{Fail} to define the statuses as variables ξ_{jt}. This is an alternative to sampling the statuses directly as random variables (as in Chapter 2).

Using this reformulation, the underlying random process is stage-wise independent, and the SDDiP can be applied. This reformulation has not been applied to supply chain disruptions that progress according to a Markovian structure.

4.3. Lagrangian cuts and implementation

To solve the model, I apply Lagrangian cuts (Zou et al. 2019). The cut coefficients, π_t^i, and v_t^i, are calculated by solving relaxations of the stage problems $\left(R_t^i\right)$. For these cuts, the linking constraints are relaxed (i.e., Constraints (44a-e)), and I solve the Lagrangian dual of the stage problems. The coefficient π_t^i is the optimal solution, π_t^*, to $\left(R_t^i\right)$, and the coefficient v_t^i is obtained as $v_t^i := \mathcal{L}_t^i\left(\pi_t^{i^*}\right)$.

$$\left(R_t^i\right): \min_{\pi_t}\left\{\mathcal{L}_t^i(\pi_t) + \pi_t^T x_{t-1}^i\right\} \quad (50)$$

Where:

$$\mathcal{L}_t^i(\pi_t) = \max d\left(q\theta_t - c^{raw}\sum_{j\in J^{API}} u_{jt} - c^{prod}\sum_{j\in J^{Line}} w_{jt}\right) - \sum_{h\in H}\sum_{j\in J^h}\left[f_h x_{jt} + c_h^{LT}\sum_{k\in\{1,\dots,L^h\}} y_{jkt}\right] - f^{Program}\phi_t + \lambda_t - \pi_t^T \tilde{z}_t \quad (51)$$

Subject to:

Constraints (14-43; 45; and 47-48)

Then the SDDiP algorithm can be applied.

It is presented in Figure 11. The parameter, Q^{im}, represents the profit in the current stage for iteration $i \in I$ and simulation sample $m \in \{1, \dots, \widetilde{M}\}$, where $\widetilde{M}$ is the number of simulations. $\dot{Q}_t^m$ is used as a record-keeping parameter for sample $m \in \{1, \dots, \widetilde{M}\}$ in stage $t \in T$. $\tilde{Q}^m$ is the total profit across all stages in sample $m \in \{1, \dots, \widetilde{M}\}$.

The model was programmed in Julia v1.2 (Bezanson et al. 2017). I used the SDDiP functionality within the SDDP.jl package to apply the algorithm and Gurobi 7.2 as the solver (Dowson and Kapelevich 2017, Gurobi Optimization LLC 2019). The analyses were conducted on a PC with 3.2GHz Intel Core i7 processor and 64GB of RAM.

There were nine stages with 10 realizations of uncertainty per stage. This is equivalent to one billion scenarios. Monetary values are scaled down by 100 to keep the matrices better conditioned. They are scaled up in the presentation of the results.

In each forward pass, a single path was sampled, and every 50 iterations, the algorithm sampled 200 iterations to generate a statistical bound. The algorithm terminated when the upper and lower bounds were not statistically different. After the algorithm terminated, I simulated

1,000 sample paths to calculate the expected profit and shortages as well as to see how the company responds to different disruption scenarios.

Algorithm 1: SDDiP Algorithm

Initialize $UB \leftarrow \infty, i \leftarrow 1$
while *Termination criteria not satisfied* **do**
  Sample path $\{U_{jt}^i \ \forall j \in J, t \in T\}$ from scenario tree $\mathcal{T}$
  # Forward step
  for $t \in T$ **do**
    if *t=1* **then**
      Solve $(P_{t=1}^i(\psi_{t=1}^i)$
      Record solution (z_1, λ_1)
    else if $t \geq 2$ **then**
      Solve $(P_t^i(z_{t-1}, \psi_t))$
      Collect solution (z_t, λ_t)
  # Backward step
  for $t \in \{|T|, ..., 2\}$ **do**
    # Generate Lagrangian cut
    for $r \in R_t$ **do**
      Solve (R_t^i) of $P_t^i(z_{t-1}^i, \psi_t^{i+1})$
      Record cut coefficients (v_t^i, π_t^i)
      Add cut to ψ_t^{i+1}
  # Update lower bound
  Solve $LB \leftarrow (P_1^i(\psi_{t=1}^{i+1}))$
  Set $i \leftarrow i + 1$
  # Calculate statistical upper bound
  if $\{i \ mod \ 50 = 0\}$ **then**
    for $m \in \{1, ..., \tilde{M}\}$ **do**
      for $t \in T$ **do**
        if $t = 1$ **then**
          Fix $x_1^m \leftarrow x_1^{i-1,m}$
          Solve $(P_{t=1}^i(\psi_1^i))$
        else if $t \geq 2$ **then**
          Solve $(P_t^i(x_{t-1}^i, \psi_t^i)$
        Record solution, z^{im}, and profit, Q^{im}
        Set $\dot{Q}_t^m \leftarrow Q^{im}$
      $\tilde{Q}^m \leftarrow \sum_{t \in T} \dot{Q}_t^m$
      # Calculate bound
      $\hat{\mu} \leftarrow \frac{1}{\tilde{M}} \sum_{m=1}^{\tilde{M}} \tilde{Q}^m; \sigma^2 \leftarrow \frac{1}{M-1} \sum_{m=1}^{M} (\tilde{Q}^m - \hat{\mu})^2; \ UB \leftarrow \hat{\mu} + z_{\alpha/2}(\frac{\hat{\sigma}}{\sqrt{\tilde{M}}})$

Figure 11. SDDiP Algorithm

5. Numerical study

A generic oncology drug, vincristine sulfate, will be used as an example of how supply chain decisions may change over time. It has been studied in the static supply chain case (Chapter 2) and using reliability modeling (Chapter 4). I will conduct analyses to evaluate the effects of reducing the lead time to add components and reducing the time to recovery. The profit results are rounded to the nearest $1,000, and percentages are rounded to the nearest 1%.

5.1. Data

The analysis applies data for the vincristine supply chain that has been previously reported in Chapter 2. It is repeated in Table 7 for reference.

The company selects from a set of candidate components. These include three API suppliers, two manufacturing plants; and five manufacturing lines. Lines are associated with particular plants, and three of the lines are in the first plant; two are in the second plant.

I apply a two-year time horizon to be consistent with typical pharmaceutical procurement contracts. The periods are three months (one quarter) which leads to nine stages overall including the initial configuration decision. The lead time to add new components is one year in the baseline analysis. The times to disruption and recovery are geometrically distributed, based on data from the FDA and the University of Utah Drug Information Service (FDA 2018a, UUDIS 2016).

Table 7. Data

	General	API Supplier	Plant	Line	Reference
Fixed costs (annual)		$34,169	$69,401	$42,200	Original source (FDA 2018b, Rudge 2012) and assumed. Calculated in Chapter 2
Fixed program cost (annual)	$9,700				(FDA 2018b, c) Calculated in Chapter 2
Lead time (annual)		Half the annual fixed costs			Assumed
Start-up costs (annual)		One-tenth the annual fixed costs			Assumed
Mean time to fail (in quarters)		69.2	112.8	34	Calculated based on FDA (2018a) and UUDIS (2016)
Failure probability (per quarter)		0.014	0.009	0.029	Calculated
Mean time to recover (in quarters)		4.8	3.2	0.32	Calculated from UUDIS (2016)
Recover probability (per quarter)		0.18	0.268	0.956	Calculated
Steady-state availability		0.929	0.968	0.971	Calculated
Lead time (in quarters)		4	4	4	
Raw material (unit)	$0.34				PharmaCompass (2018) and procurement representatives
Production cost (unit)	$2.22				Calculated in Chapter 2
Price (unit)	$5.55				(IBM Micromedex 2018)
Demand (annual units)	90,000				CMS (2018a, b) and National Cancer Institute (2018) Calculated in Chapter 2
Objective function bound	$2dq\|T\|$				Calculated

5.2. Numerical results

Baseline (1 year lead time)

In the baseline analysis, the company decides its initial configuration and has a one year lead time to add additional components. The expected profit over the time horizon is $184,000 (corresponding annual profit is $92,000) and expected shortages are 13%. The distributions are shown in Figure 12. The company is profitable most of the time; in 96% of the simulated runs,

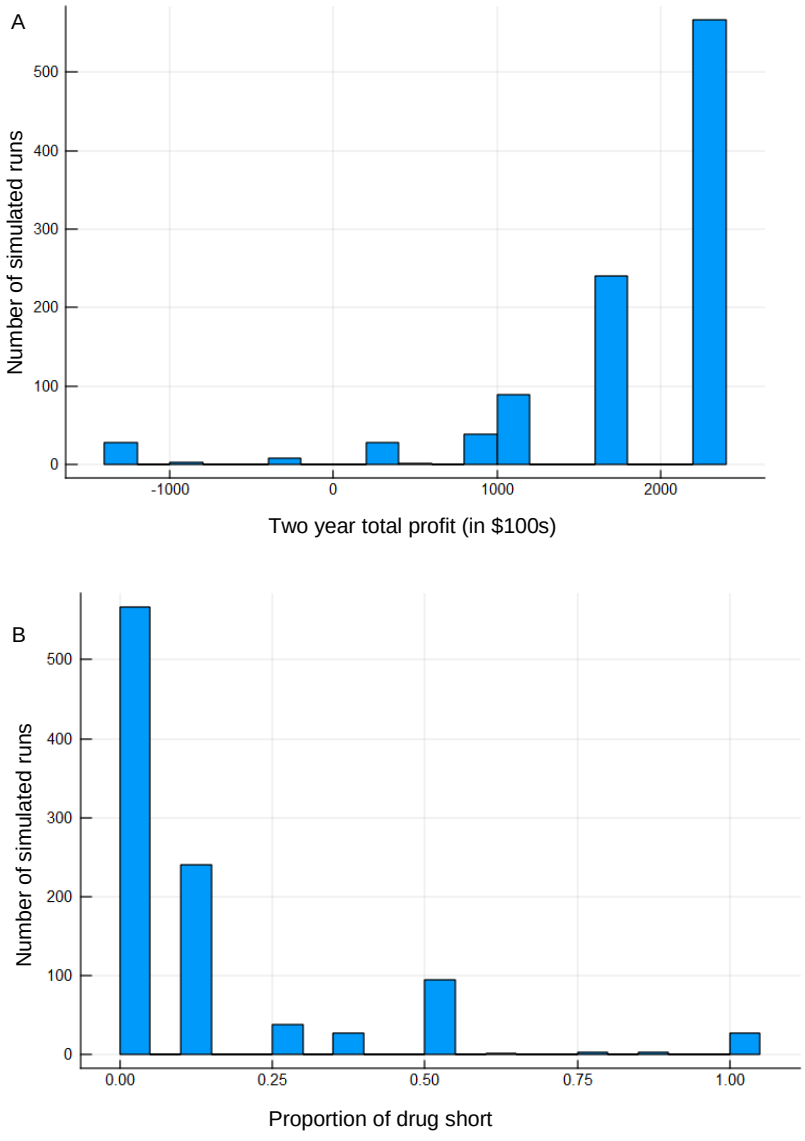

Figure 12. Total profit and shortages (1 year lead time)

the company makes money. In 57% of the runs, there is not a shortage, though in 19% of the runs, the drug was short for at least two quarters.

The company remains in the market in 88% of the simulated runs. Among discontinuations (i.e., when the company leaves the market), 9% occur in the 6th stage. Recall that the model enforces that companies may not leave the market in the last 2 stages; at this point they either decide to remain in the market until the end of the horizon or discontinue. The distribution of decisions over time is shown in Figure 13 panel B. The shading represents the percent of simulations. The lightest shading is the 0-100th percentiles. For panel B, observe that the company was in the market for all of the simulation runs until stage 5. At stage 5, the company leaves the market in 3% of the simulated runs. At stage 6, the company leaves the market in 9% of the simulated runs. This is represented with the medium shading, reflecting the proportion of runs within the 10-90th percentiles.

Similarly, we observe the proportion of shortages over time in panel A. Shortages are higher in the latter stages, which is shown via darker shades with lower demand met.

In none of the simulated runs did the company add components.

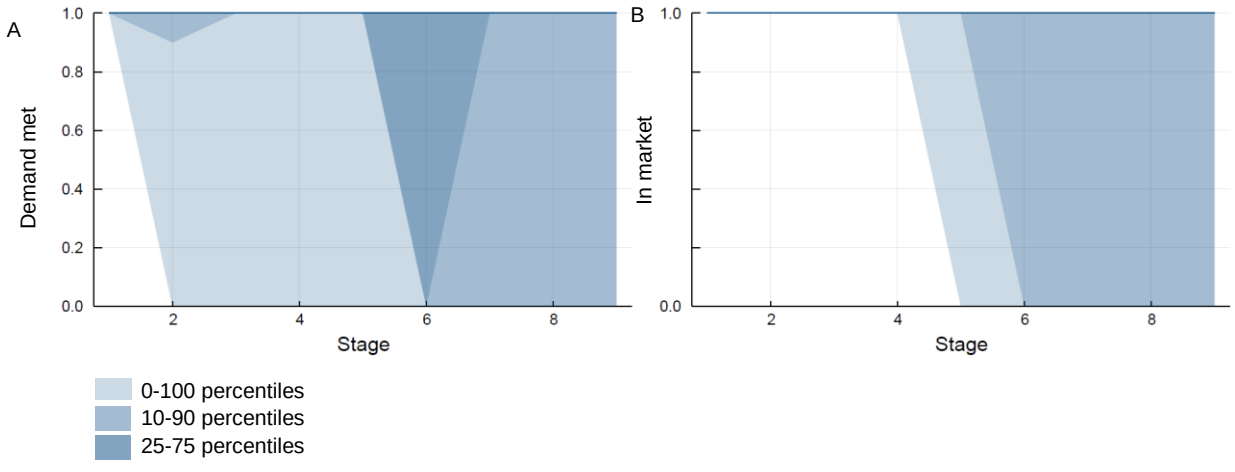

Figure 13. Demand met and proportion in market over time (1 year lead time)

Next, we will look at the results for specific sample paths. The figures will be presented with four panels. The supply chain design is presented in panel A, in which the black elements represent that a component is in the configuration in a given stage. Panel B shows the status of each of the candidate components for each stage; green represents available, and white is unavailable. Panels C and D are the profit and shortages in each stage.

In many of the sample paths, there are no disruptions. An example is given Figure 14. The company selects a lean supply chain and maintains the same configuration for the entire time horizon. The profit is negative in the initial stage because of the start-up costs, and positive once the company starts meeting demand (stages 2-9). The total profit is $234,000, and all demand is met.

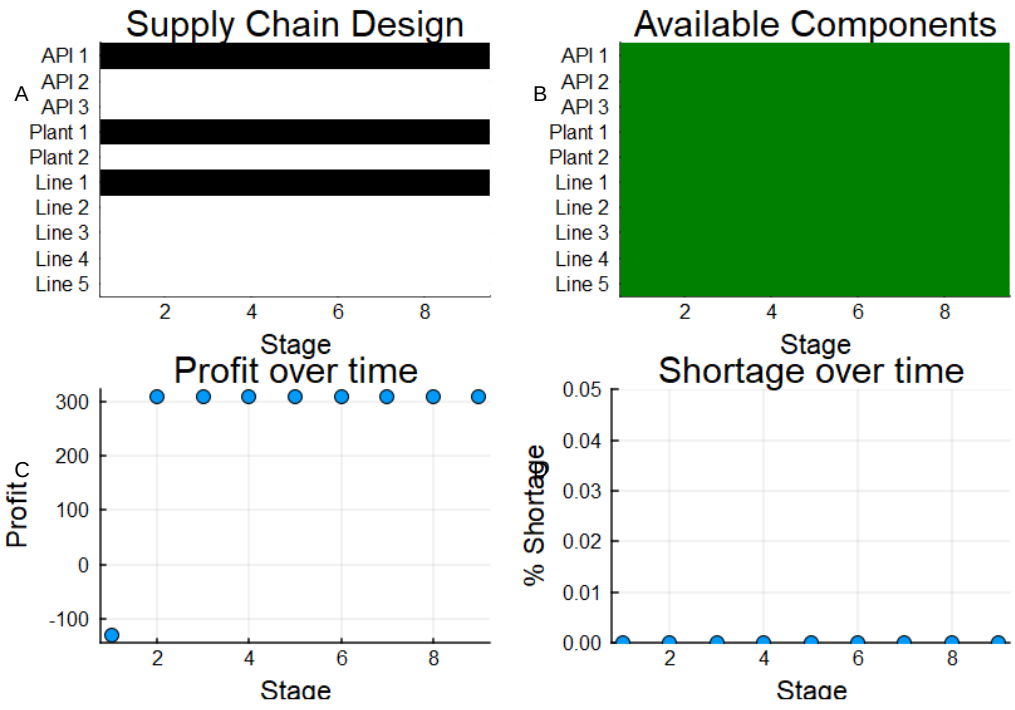

Figure 14. Results for sample path without a disruption (1 year lead time)

In another sample path presented in Figure 15, the selected plant is disrupted for a single stage. The company does not make changes to its supply chain. Rather it waits to see if the component continues to be disrupted in future stages before it makes a change (e.g., adding a

second plant or leaving the market). During the disruption, the company loses money (i.e., paying the fixed supply chain costs without accruing revenue), and there is a shortage. The total profit is $166,000, and the drug is short 12.5% of the time.

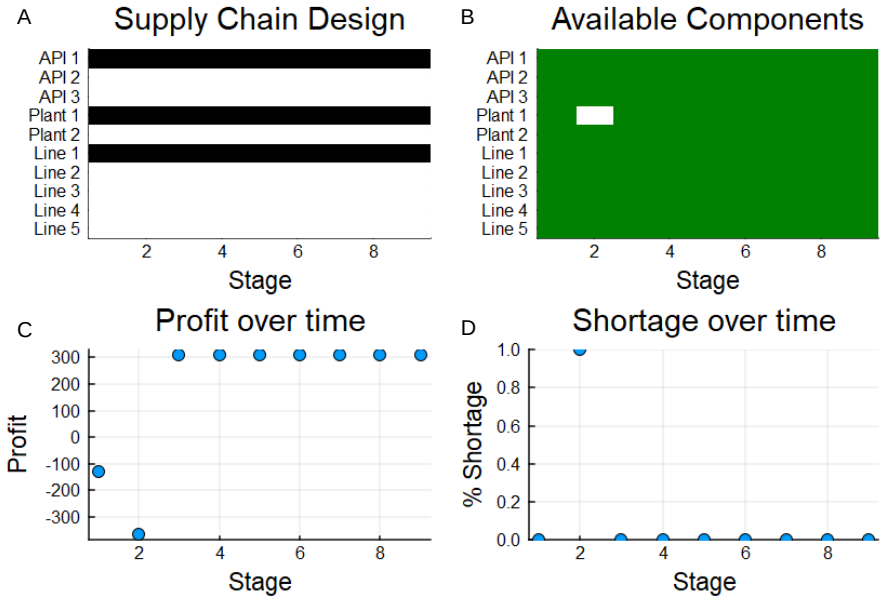

Figure 15. Results for sample path with a brief disruption (1 year lead time)

Six month lead time

This analysis will apply a lead time of six months (2 stages) for a component to be added. The results are very similar to the case in which there is one year of lead time. The profit differences are likely based on more variation in the underlying uncertainty than the change in lead time itself. The expected shortages are 13%, and the expected total profit is $182,000. The distribution of shortages and market status over time are shown in Figure 16. In a handful of the simulated runs (0.6%, n = 6), the company adds a second supplier. These correspond to a small increase (0.6%) in the proportion of sample paths where the company remains in the market. The company does not add a plant in any of the runs.

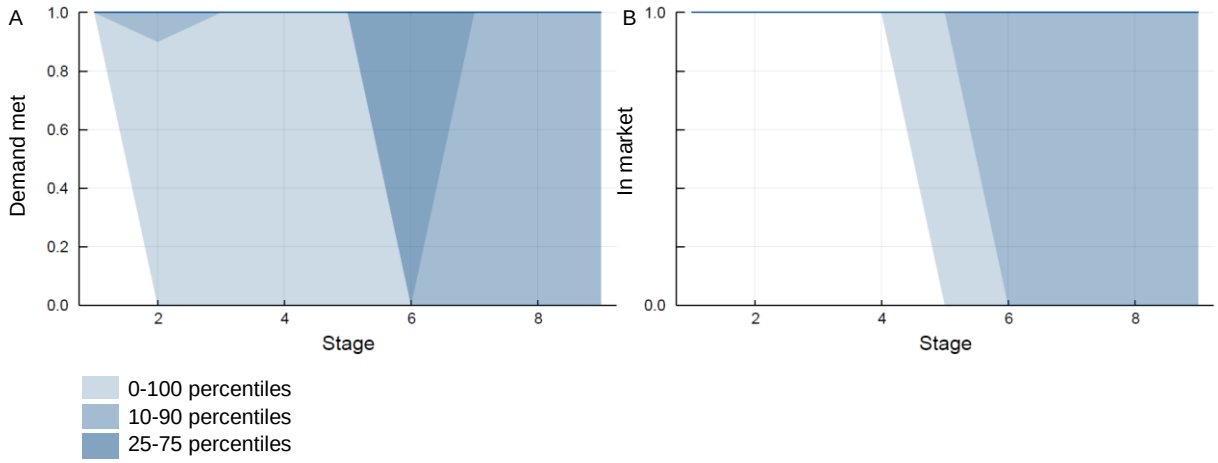

Figure 16. Demand met and proportion in market over time (6 month lead time)

One path where the company adds a second supplier is when there is a supplier disruption in the 6th stage (Figure 17). Panel A represents which components are in the configuration. Panel B is the availability of each of the components over time. The company selects to bring up an additional supplier when the disruption occurs; it is operational in the final 2 stages. The annual profit is $37,000 with shortages of 25%. Had the company instead decided to leave the market in the 6th stage, shortages would have been 50%.

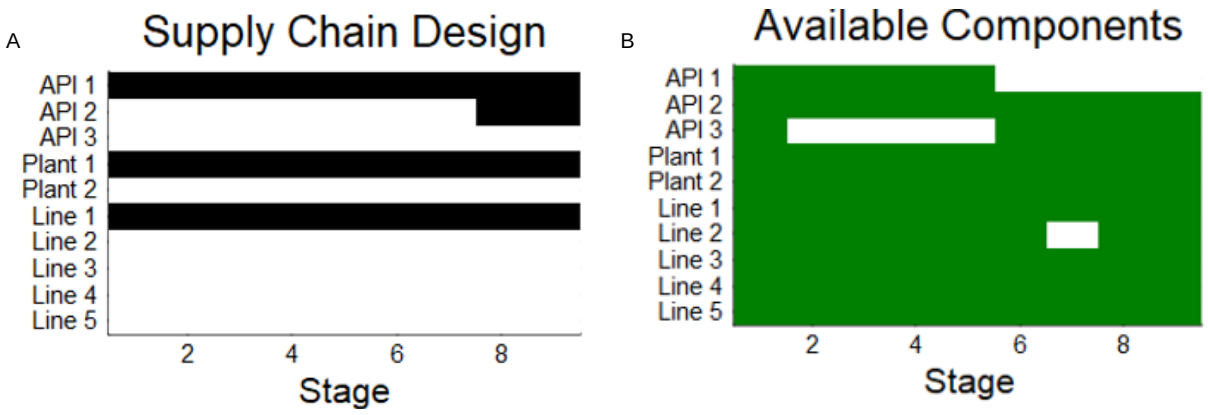

Figure 17. Disruption towards end of horizon (6 months of lead time)

Three month lead time

With three months of lead time, there begins to be a decrease in shortages. Expected shortages decline to 9% (vs. 13% with longer lead times), and the expected total profit is slightly higher at $186,000. The distribution of shortages and profit in the simulated paths are shown in Figure 18. In 88% of the runs, shortages occur for a single quarter or not at all.

Prominently, the company decides to leave the market less often. Comparing the outcomes in Figure 16 (six month lead time) with Figure 19 (three month lead time), we see effects of the reduced lead time in the later stages. Demand is met more often in the latter stages (A panels), and the company decides to leave the market less often (B panels). This is

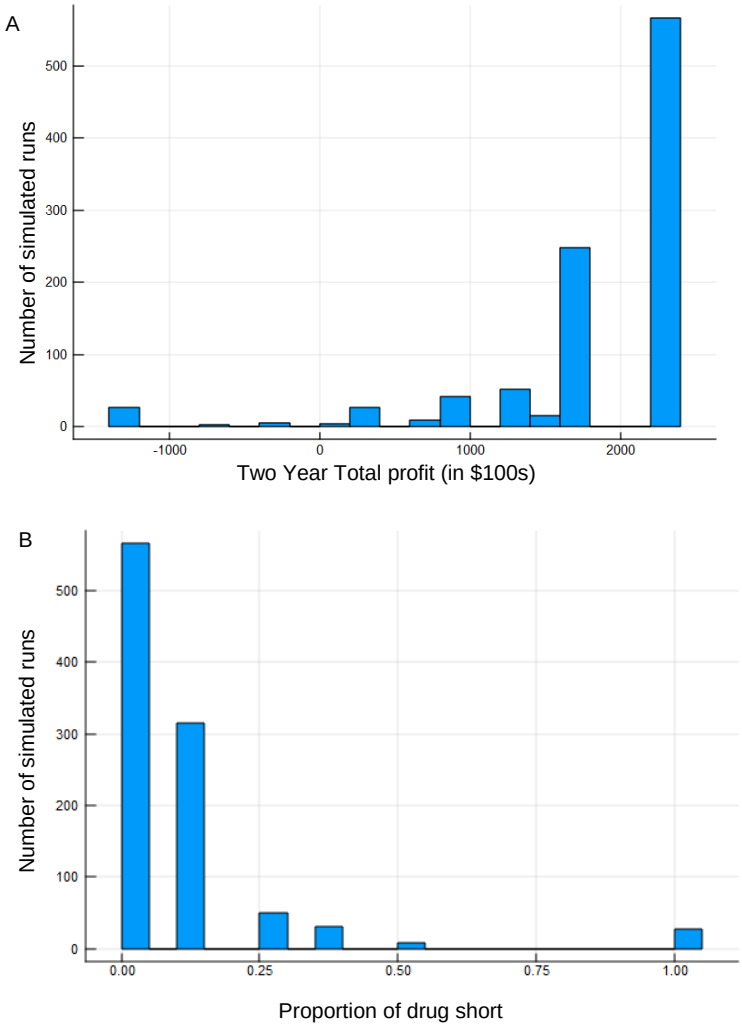

Figure 18. Distribution of total profit and shortages by simulated path (3 month lead time)

represented with lighter shadings in the latter stages, reflecting fewer simulated runs where the company leaves the market.

With a lead time of three months, the company leaves in only 3% of the simulated paths (vs. 12% with longer lead times).

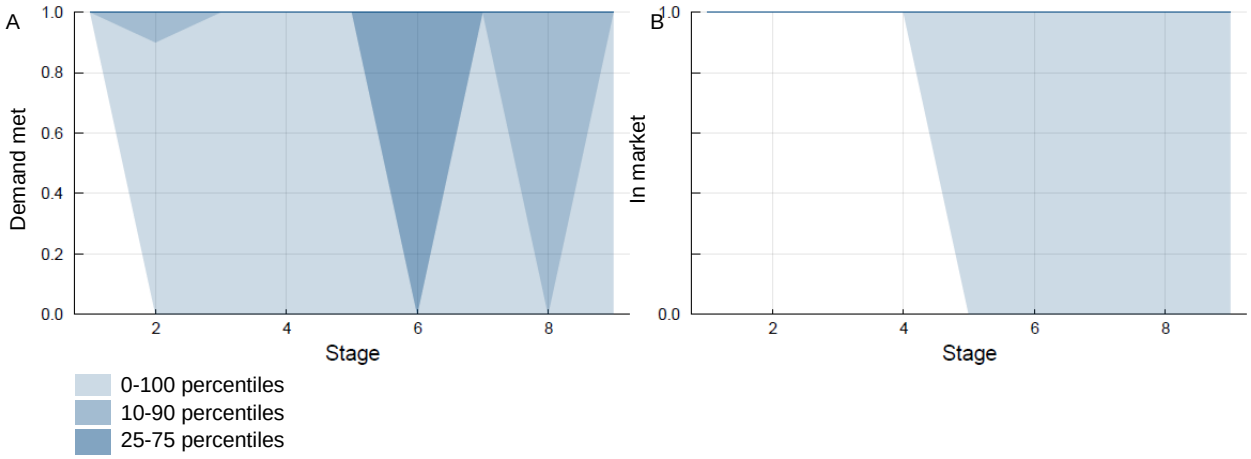

Figure 19. Demand met and proportion in the market over time (3 month lead time)

The company adds an API supplier in 10% of the simulated paths (n = 100). One example is presented in Figure 20. The total profit is $144,000, and the drug is short 12.5% of the time. In the 6th stage, a disruption occurs to the selected supplier occurs (API 1), and the drug is unavailable. The company begins the process of bringing up a second supplier (API 2), and it is available in the next stage because of the three month lead time. After the first supplier recovers in the final stage, the company removes the second supplier and returns to a lean supply chain. In contrast, the company does not select an additional plant in any of the simulated paths.

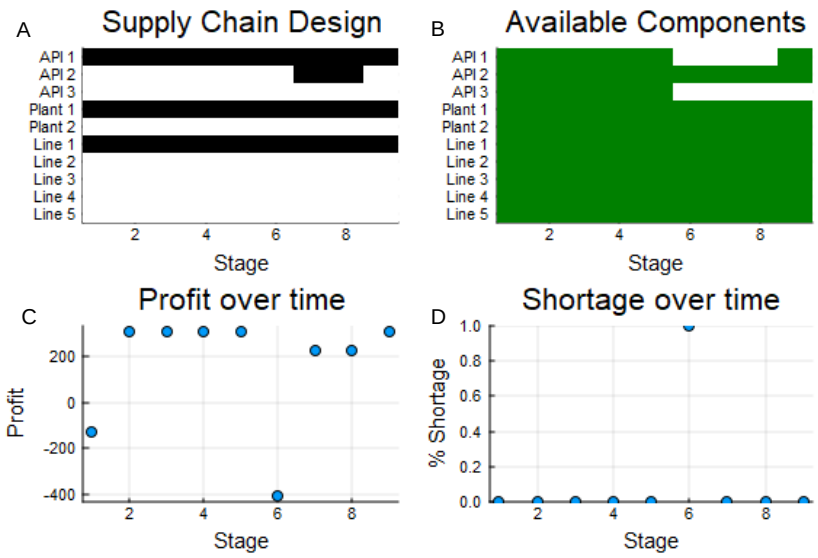

Figure 20. Results of supplier disruption (3 month lead time)

Halve the mean time to recover

In the final analysis, I evaluate an alternative intervention: reducing the recovery time. The previous analyses focused on reducing the time to add components to the supply chain; in this, the focus is on the components that have already been selected. I apply a lead time of one year and reduce the mean time to recover by half for each of the components.

The expected shortages are 9%, and the expected total profit is $202,000. In 88% of the simulated paths, shortages last for at most a quarter. At the end of the horizon, in 90% of the paths, the company is still in the market. Figure 22 presents the distributions of expected profit and shortages across sample paths. The distributions of market status and shortage outcomes are available by stage in Figure 21.

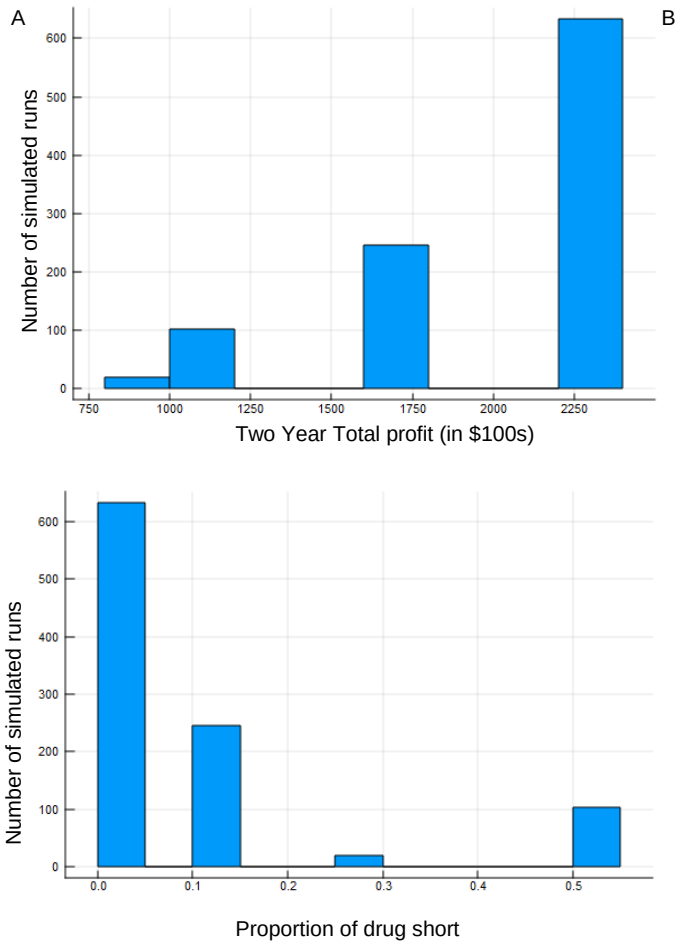

Figure 22. Total profit and shortages (1 year lead time; half mean time to recover

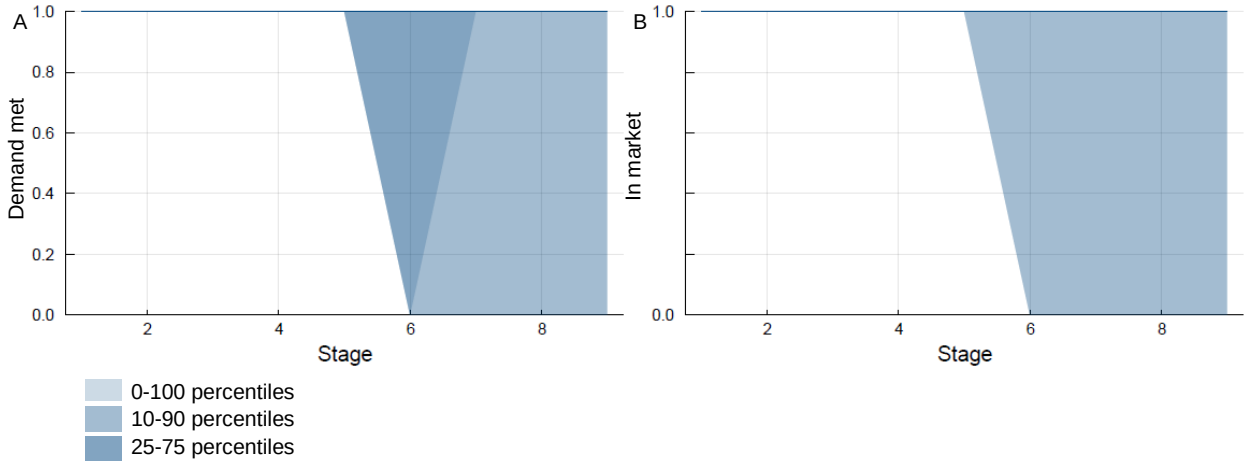

Figure 21. Proportion of demand met and proportion in the market (1 year lead time; half mean time to recover)

As compared to the baseline, expected shortages decrease (13% to 9%). This is comparable to the three month lead time analysis. The expected profit is higher than in either the baseline or the lead time analyses. The benefits compared to the baseline are intuitive; as components are available more often, the company provides the drug more often and makes more revenue. The profit is higher compared to the three month lead time analysis because the company is not incurring the costs of additional components.

Table 8 provides a summary of the main results from each analysis. The expected total profit does not change substantially among the lead time analyses. It increases as the mean time to recover is halved. Expected shortages are lowest when the lead time is 3 months and when the mean time to recover is halved. The 3 month lead time analysis has the highest proportion of simulated runs where the company remains in the market (97%).

Table 8. Summary of analyses

Analysis	**Expected total (2 year) profit**	**Expected shortages**	**Percent in the market at the end of the horizon**
1 year lead time	$184,000	13%	88%
6 month lead time	$182,000	13%	88%
3 month lead time	$186,000	9%	97%
Half mean time to recover	$202,000	9%	90%

6. Discussion

In this chapter, I present a new model of a dynamic supply chain configuration with components that may be disrupted and recover. I apply the model in the pharmaceutical context and consider the effects of dynamic configurations on drug shortages. The analyses suggest that

decreasing the lead time to bring new components up or the time to recover can reduce shortages.

An earlier chapter considered the static design problem under disruption (Chapter 2). In the baseline analysis of the dynamic model, with a lead time of one year to bring up new components, shortages are greater than the static case (13% vs. 11%). This is initially counterintuitive as the dynamic model has the ability to adapt to disruptions when they occur. However, the objective is to maximize the expected profit, not to minimize shortages, and when a disruption occurs, a dynamic decision-maker can choose to leave the market and avoid paying the fixed costs. In the static model, a company that experienced a disruption would remain in the market. The higher percentage of shortages at the baseline for the dynamic model is a result of these differences between the models.

In all of the analyses, the company is more likely to leave the market later in the time horizon. The earliest a company leaves the market is in the 5^{th} stage. In the quick-recovery analysis, the company does not leave in any of the sample paths until the 6^{th} stage. The proportion of paths which the company leaves varies depending on the lead time. As the lead time to bring on new components decreases, the proportion of paths where the company discontinues production decreases, from 12% with one year of lead time to 3% with three months of lead time.

As the lead time decreases, shortages may also decrease as well. There is no difference between the one year and six month lead time cases (both exhibit 13% shortages), though decreasing the lead time to three months reduces expected shortages to 9%. Halving the mean time to recover also leads to expected shortages of 9%. This suggests that either could be effective at reducing shortages.

The lead time results suggest that it is worthwhile to consider ways to reduce the barriers to adding capacity. This could include reducing regulatory review times and pre-qualifying components. There is some evidence that when the FDA has expedited reviews for new facilities that produce short drugs, shortages have been reduced (GAO 2016). There could also be incentives for dual-sourcing. For the quick-recovery case, shorter review times would also be beneficial, and companies could explore ways to improve response to disruptions. In practice, a company could also hold inventory to meet demand when capacity is unavailable, though, this is not frequently done for generic injectables.

However, in all analyses, shortages continue to remain high (9-13% across the analyses). This suggests that reducing the lead time or reducing the time to recovery may not be enough to eliminate the drug shortage crisis. It may be necessary to implement other types of interventions such as legislation or contractual changes (Jia and Zhao (2017), Chapter 2).

When a disruption occurs, a company may choose to leave the market. This occurred in the case of the oncology drug Doxil (Palmer 2013). The plant was disrupted, and the company decided to not re-open the plant. In these analyses, some of the shortages occur because the company chooses to leave the market. This highlights the value of considering discontinuation decisions in disruption models. In disruption models, it is often assumed a company aims to return to the initial state (Hopp and Yin 2006), but it may not be optimal to do so, particularly when the profit margins are tight.

The analyses used a low-profit-margin drug as a case example. It is optimal to have a vulnerable, lean supply chain in the baseline analysis, which may not be the context for other types of products. Nonetheless, the modeling framework is applicable to other settings. Many areas have embraced lean supply chains, and disruptions can have major ramifications on a

company's ability to supply its products. In practice, companies may add components to an existing supply chain after a disruption, though this has received limited study in the literature. Nearly always, capacity expansion problems assume stochastic demand (e.g., Yu et al. 2018).

In this chapter, I allow the company to adjust which components are selected as a mitigation strategy and apply the SDDiP algorithm. It is well-suited to strategic supply chain decisions as it is often natural to model component selections using binary variables. Recent work has applied SDDiP to the facility location problem with stochastic demand (Yu et al. 2018). This chapter study provides evidence of its utility for disruption modeling.

Often disruption models will sample disruptions from a Bernoulli distribution (Snyder et al. 2016). This would work well with the SDDiP algorithm because these random variables are naturally stage-wise independent. With the multiple distributions for the times to disruption and recovery, the uncertainty is initially stage-wise dependent. I reformulate the random variables and constraints to apply stage-wise independence. The change is in the spirit of sampling random errors of an autoregressive process instead of the process itself (Shapiro 2011), though in the case of geometric distributions, I sample the probability values instead. In both contexts, the random process is applied in the constraints. The reformulation provides a new framework to apply Markovian uncertainty in SDDiP, without an exponential expansion of the state space that quickly becomes intractable (Philpott and De Matos 2012).

The analysis is limited in that a single drug is considered as a case example. It is not representative of all pharmaceutical products, and results should not be applied without further study of the effects on other products. In addition, because some data were estimated or assumed, the results should be interpreted as indicating the magnitude of effects rather than precise estimates.

7. Conclusions

Drug shortages continue in the US, and better supply chain models are needed to represent the disruption and recovery dynamics of the pharmaceutical industry. I address the case where the configuration of the supply chain may change in response to a disruption's occurrence. The endogenous component decisions allow the company to add capacity or discontinue production as availability changes. I reformate the Markov process for availability to implement stage-wise independent uncertainty and highlight the use of the SDDiP algorithm for supply chain design problems. The analyses suggest that shortening the time to add capacity or reducing the time to recover disrupted capacity may reduce shortages. Future work could consider competition between companies or incorporating endogenous demand.

CHAPTER IV

Pharmaceutical Supply Chain Reliability: A New Model and Analysis of Drug Shortages

1. Introduction

The pharmaceutical industry is subject to disruptions that range from natural disasters to supplier failures. Supply chain disruptions are a major cause of drug shortages (UUDIS 2016) and improving risk assessment could help reduce them. Supply chains of generic injectable drugs are particularly vulnerable. They often lack redundancy (Woodcock and Wosinska 2013), and disruptions can quickly lead to shortages. These shortages have large effects on patient care and health system costs (Tucker et al., 2020a). While it may be more profitable for companies to maintain the unreliable supply chains that can lead to shortages (Chapter 2), some have reported that they manufacture drugs with limited profitability, in part, because of their need in patient care (GAO 2011).

I believe that it is important to provide drugs for patients who need them. My other work studies how incentives could increase the ability of companies to produce a resilient and profitable drug supply (Chapter 2) and how improving disruption response could reduce shortages (Chapter 3). These have been modeled using stochastic programs, and the focus has been *optimizing* the configuration. Simpler models can be developed for the *evaluation* step. In this chapter, I study given supply chain configurations. I develop simple, closed-form expressions for pharmaceutical supply chain reliability.

For a given supply chain configuration, the model reports the expected reliability, average shortage length (downtime), and average time between shortages (uptime). I consider a generic oncology drug as a case example and study the effects of different configurations on drug shortages. I also evaluate the profitability of different configurations under varying prices.

This model could be used by regulators or companies to quickly evaluate the reliability of a pharmaceutical supply chain configuration. These evaluations could occur during risk assessments or in external evaluations. The model could also be extended to be used for other supply chain structures or in other industries.

2. Literature review

There is a wide literature on drug shortages. A recent review surveyed 430 papers that have been published or disseminated since 2001 (Tucker et al., 2020a). Much of the research focused on the health effects of shortages or on what causes them (e.g., Vail et al. 2017). Far fewer were quantitative studies of supply chains, despite the strong connection between supply chain management and shortages.

Among those that have considered supply chains, some have studied contractual and legislative policies to reduce shortages (Jia and Zhao 2017, Chapter 2). Others considered human behavior in supply chain decisions (Doroudi et al., 2018) and how competition affected spare capacity decisions (Kim and Scott Morton 2015). Downstream, there has been work to optimize hospital inventory in response to shortages (Saedi et al. 2016) and recalls (Azghandi et al. 2018) as well as to stockpile pediatric vaccines (Jacobson et al. 2006). There are not models to evaluate the risk of shortages for pharmaceutical supply chains.

More broadly, supply chain resiliency is an active field. Recent reviews (Pires Ribeiro and Barbosa-Povoa 2018, Snyder et al. 2016) noted the wide range of research. To help

companies consider resiliency, researchers have developed overarching frameworks. Kleindorfer and Saad (2005) focused on disruption management and reported steps to understand and mitigate risks. These steps may include considering the categories and sources of disturbance (Svensson 2000) and considering vulnerabilities and capabilities (Pettit et al. 2010). Asbjørnslett (2009) gave a seven-step process that included classification of vulnerability factors, evaluation of vulnerability scenarios, and mitigation. Two phases – understanding and mitigation – resonate throughout these works. To be resilient a company needs to understand its supply chain and its susceptibility to disruptions and ensure satisfactory steps are taken to address risk factors.

An important step in the understanding phase is to conduct a vulnerability assessment. In this, the company evaluates its exposure to risk. These could be global suppliers, dependence on particular suppliers or customers, or single-sourcing (Wagner and Bode 2006). One way to measure risk is to use an index. This can evaluate the current supply chain as well as mitigation options.

A popular approach for indices is to apply graph theory. This has led to the Supply Chain Resilience Index (SCRI) (Soni et al. 2014) and the Supply Chain Vulnerability Index (SCVI) (Wagner and Neshat 2010), with a follow-up empirical study (Wagner and Neshat 2012). The effects of disruptions on an automotive supply chain was studied using a risk-exposure index (Simchi-Levi et al. 2015).

Other methods have also been used. A time-series model (an auto-regressive integrated moving average model [ARIMA]) was applied to generate vulnerability indicators (Sakli et al. 2014). Their indicators included delays, inventory levels, and over-cost. A simulation analysis found that delays were an effective metric for measuring impact on the supply chain (Vilko and

Hallikas 2012). The model was parameterized using data gathered from interviews of many supply chain stakeholders. Copulas have been used to consider correlations (Jia and Cui 2012), and fuzzy methods have been used to incorporate subjective inputs (Liu et al. 2016, Samvedi et al. 2013).

There are also non-index methods to understand a supply chain's susceptibility to risk. In a fault-tree approach, logic gates are used to represent a supply chain. The analysis showed the risk of delay (Sherwin et al. 2016) and can be overlaid with an optimization model to produce effective mitigation strategies (Sherwin et al. 2020). To consider interactions between components, dynamic fault-trees have also been used (Lei and MacKenzie 2019).

The propagation of disruptions has been analyzed using network modeling (Wu et al. 2007). Disruption severity has also been considered in a conceptual study (Craighead and Blackhurst 2007) and using network simulation (Adenso-Diaz et al. 2012).

Some assessment methods are based on reliability modeling. These include an early study of contingency logistics systems (Thomas 2002) and more recent work on interdependent suppliers (Hagspiel 2018). In a paper related to this work, Ha, Jun and Ok (2018) present reliability functions for different supply chain configurations. This chapter considers different configurations and analyzes them in the context of drug shortages.

Within the pharmaceutical industry, there is a need for models to evaluate supply chain reliability. This is particularly important when the company is the sole manufacturer of the drug. Reliability metrics could affect strategic decisions as companies design or re-design their supply chains. In addition, if supply chain information becomes available to regulators or the public (as proposed by ASHP (2018b)), external stakeholders could also use the model to assess strategic shortage risks.

The contributions of this chapter are:

- The development of new reliability functions for pharmaceutical supply chains subject to disruption
- A case study of generic injectable oncology drugs to consider the effects of reliability on drug shortages
- A pricing analysis, including break-even drug prices and optimality thresholds

The rest of the chapter will proceed as follows. Section 3 presents the supply chain reliability formulation. In Section 4, I conduct a numerical study on drug shortages, and in Section 5, I present the pricing analysis. I discuss the results in Section 6 and conclude in Section 7.

3. Supply chain reliability (SCR)

3.1. Overview

Within pharmaceutical supply chains, disruptions to the API suppliers, manufacturing plants, and manufacturing lines most often cause shortages (GAO 2014, UUDIS 2016). The focus will be on these three types of components. The goal is to produce closed-form equations of the reliability of a pharmaceutical supply chain. The proposed Supply Chain Reliability (SCR) model can be used to calculate the reliability and expected shortages of echelons of a pharmaceutical supply chain where disruptions may lead to shortages.

Given the number of API suppliers, plants, and lines in the supply chain of a single drug, it will output: the reliability of the supply chain, the corresponding expected shortages, the average uptime (time between shortages), and the average downtime (length of a shortage). Reliability is defined as the probability of a supply chain being able to meet demand. Each of the components are either available or disrupted.

In the supply chain configuration, there are two stages: the API suppliers and the plant-line combinations. The lines are associated with specific plants, but the API suppliers can send materials to any plant (as in Chapter 2). The stages operate in series; at least one supplier and one plant-line combination must be working for the drug to be produced. Within the stages, the components operate in parallel; only one component is needed for the stage to be considered available. That is, only one API supplier needs to be available for the API to be supplied, and only one plant-line combination needs to be available to produce the finished form of the drug. Each component has adequate capacity to supply call demand.

Each component fails or recovers independently of the other components. I consider the overall risk of disruption rather than specific types of disruptions. The rates of component failure and recovery can be based on any continuous distribution and vary by the type of component. For example, lines can have different recovery rates than suppliers.

The supply chain as a whole is able to supply the drug if there is at least one supplier available and at least one plant-line combination available. In this case, the company is able to meet demand, called "system availability." If the company is not able to meet demand, this is designated "system disruption."

3.2. Sets and notation

The notation for the sets, parameters, and outcomes is presented in Figure 23. It follows the notation in Ross (2014). The bold terms are vectors, and the non-bold terms are scalars.

Sets

H	Set of echelons, $H = \{API, p, l\}$
N	Set of all components
N^h	Set of components in echelon $h \in H$, where $N^h \subset N$

Parameters

z^{API}	Number of API suppliers selected
z^p	Number of plants selected
z^l	Number of lines selected in each plant
$\boldsymbol{z}$	Vector of components per echelon, i.e., z^{API}, z^p, z^l
$\boldsymbol{\mu}$	Recovery rate, a vector comprised of the recovery rates for each echelon, i.e., μ^{API}, μ^p, μ^l
$\boldsymbol{\lambda}$	Disruption rate, a vector comprised of the disruption rates for each echelon, i.e., $\lambda^{API}, \lambda^p, \lambda^l$

Outcomes

s	Expected shortages
$\overline{U}$	Average system uptime
$\overline{D}$	Average system downtime

Figure 23. Notation for sets, parameters, and outcomes

The structure function, $\phi(\boldsymbol{X}, \boldsymbol{z})$, reports whether demand can be met given the current status of the supply chain components. The notation for the structure function and probabilities are summarized in Figure 24. The reliability function, $r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right)$, is the probability that demand can be met for a particular configuration, i.e., the expectation of the structure function. The probability the system is available given that a particular component $n \in N$ is available is: $r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 1\right)$.

I also define reliability functions for specific stages of the supply chain. The probability that raw materials are available is $r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}\right)$, and the probability that finished goods can be produced is $r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}\right)$. These represent the probabilities at least one supplier is available and a least one plant-line combination is available, respectively. Finally, for a given supply chain $\boldsymbol{z}$, the probability that the failure of a particular component $n \in N$ in echelon $h \in H$ leads to a system failure is $\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}\right)$.

$\phi(\boldsymbol{X}, \boldsymbol{z}) = \begin{cases} 1 & \text{if supply chain } \mathbf{z} \text{ is available if the components have status } \boldsymbol{X} \\ 0 & \text{otherwise} \end{cases}$

$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}\right) = P\{\phi(\boldsymbol{X}, \boldsymbol{z}) = 1\} = E[\phi(\boldsymbol{X}, \boldsymbol{z})]$ Reliability function given system $\boldsymbol{z}$

$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}, X_n = 1\right)$	probability the system $\boldsymbol{z}$ is available given component $n \in N$ is working
$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}, X_n = 0\right)$	probability the system $\boldsymbol{z}$ is available given component $n \in N$ is not working
$r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| z^{API}\right)$	probability at least one supplier is working given z^{API} suppliers
$r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| z^p, z^l\right)$	probability at least one plant-line combination is working given z^p plants and z^l lines in each plant
$\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \middle| \boldsymbol{z}\right)$	probability that the failure of component $n \in N$ in echelon $h \in H$ causes a system failure

Figure 24. Additional notation

3.3. Model

The model is three equations: the expected shortages (s), the average uptime ($\bar{U}$), and the average downtime ($\bar{D}$). These are exact, closed-form equations given the failure rates (λ), recovery rates (μ), and supply chain configuration (N^h).

Reliability and expected shortages

Reliability is the probability the system is able to produce the drug in the long-run. The equation is presented in (1).

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) \quad (1)$$

It is the product of the probability that raw materials can be ordered and the probability that the finished form can be produced.

The probability that raw materials can be ordered is the probability at least one supplier can supply them, equation (2). It is one minus the probability that all of the suppliers are unavailable.

$$r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) = 1 - \left(\frac{\lambda^{API}}{\mu^{API}+\lambda^{API}}\right)^{z^{API}} \quad (2)$$

The probability that the finished form can be produced is the probability that at least one plant-line combination is available, equation (3). It is one minus the probability that in each plant: either the plant is unavailable or all of lines in a working plant are unavailable.

$$r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) = 1 - \left(\left(\frac{\lambda^{p}}{\mu^{p}+\lambda^{p}}\right) + \left(\frac{\mu^{p}}{\mu^{p}+\lambda^{p}}\right)\left(\frac{\lambda^{l}}{\mu^{l}+\lambda^{l}}\right)^{z^{l}}\right)^{z^{p}} \quad (3)$$

It follows that expected shortages are one minus the reliability, equation (4).

$$s = 1 - r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right) \quad (4)$$

Average uptime and downtime

The average uptime is given in equation (5) and the average downtime is given in equation (6). These represent the average time between shortages and the average time to recover, respectively.

$$\overline{U} = \frac{r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right) r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)}{z^{API}\left[\frac{\lambda^{API}\mu^{API}}{\mu^{API}+\lambda^{API}}\right]\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|z\right)+z^{p}\left[\frac{\lambda^{p}\mu^{p}}{\mu^{p}+\lambda^{p}}\tilde{r}^{Plant}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)+z^{l}\left[\frac{\lambda^{l}\mu^{l}}{\mu^{l}+\lambda^{l}}\right]\tilde{r}^{Line}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\right]} \tag{5}$$

$$\overline{D} = \frac{\overline{U}}{r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right) r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)} - \overline{U} \tag{6}$$

Derivations

Given a system of independently available and disrupted components, a general expression of the average uptime is shown in equation (7) (Ross 2014). The only restriction is that the availability and distribution distributions be continuous. It is the probability the component is available divided by the rate individual components lead to failure.

$$\overline{U} = \frac{r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)}{\sum_{n\in N}\frac{\lambda_n\mu_n}{\lambda_n+\mu_n}\left[r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|\boldsymbol{z},X_n=1\right)-r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|\boldsymbol{z},X_n=0\right)\right]} \tag{7}$$

I define the probability a component failure causes system failure, i.e., $r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z},X_n=1\right)-r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z},X_n=0\right),\forall n\in N^h,h\in H$ as $\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)$.

Then the average uptime can be written as in equation (8).

$$\overline{U} = \frac{r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)}{\sum_{h\in H}\sum_{n\in N^h}\frac{\lambda_n\mu_n}{\lambda_n+\mu_n}\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)} \tag{8}$$

Substituting the average reliability produces equation (9).

$$\overline{U} = \frac{r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right) r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)}{\sum_{h\in H}\sum_{n\in N^h}\frac{\lambda_n\mu_n}{\lambda_n+\mu_n}\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\middle|\boldsymbol{z}\right)} \tag{9}$$

Finally, the probabilities that component failure causes system failure (derived below) are incorporated, and the equation simplifies to Equation (5).

The downtime is similarly derived using the general Equation (10) from Ross (2014).

$$\overline{D} = \frac{\overline{U}\left[1-r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\right]}{r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)} \tag{10}$$

Component failure causes system failure

The time to fail is based on i) the probability a given component causes the system to fail and ii) the probability of the system is available at any given time.

The probability that a given component failure causes system failure is calculated as the probability the system is available when the component is working minus the probability the system is available when the component is not working. It is given in Equation (11).

$$\tilde{r}_n^h\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right) = r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 1\right) - r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 0\right), \forall n \in N^h, h \in H \tag{11}$$

The probability that a component failure leads to a system failure differs by the type of component. Equations (12), (13), and (14) present the probabilities for API suppliers, manufacturing plants, and manufacturing lines, respectively.

$$\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right) = r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right)\left(\frac{\lambda^{API}}{\mu^{API}+\lambda^{API}}\right)^{z^A-1} \tag{12}$$

$$\tilde{r}^{Plant}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right)\left(1-\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1} \tag{13}$$

$$\tilde{r}^{Line}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right)\left(\frac{\mu^{p}}{\mu^{p}+\lambda^{p}}\right)\left(\frac{\lambda^{l}}{\mu^{l}+\lambda^{l}}\right)^{z^{l}-1}\left(\frac{\lambda^{p}}{\mu^{p}+\lambda^{p}}+\frac{\mu^{p}}{\mu^{p}+\lambda^{p}}\left(\frac{\lambda^{l}}{\mu^{l}+\lambda^{l}}\right)^{z^{l}}\right)^{z^{p}-1} \tag{14}$$

API suppliers

First, I will consider the derivation of $\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)$ for API suppliers. Equation (15) was previously defined as the difference between the probabilities that the system is working when the component is available and when the component is not.

$$\tilde{r}_{n}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z}\right)=r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z},X_{n}=1\right)-r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\,\middle|\,\boldsymbol{z},X_{n}=0\right),\forall n\in N^{API} \tag{15}$$

If there is a working supplier, the system is available if a plant-line combination is available (Equation (16)).

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_{n}=1\right)=r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\quad\forall n\in N^{API} \tag{16}$$

If the given API supplier $n\in N^{API}$ is not available, there needs to be at least one other available supplier and a working plant-line combination for the system to be available. The probability this occurs is the probability at least one other supplier is available multiplied by the probability a plant-line combination is available (Equation (*17*)). The probability another supplier is available is one minus the probability all other suppliers are unavailable.

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_{n}=0\right)=\left[1-\left(\frac{\lambda^{API}}{\mu^{API}+\lambda^{API}}\right)^{z^{A}-1}\right]r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\ \forall n\in N^{API} \tag{17}$$

Then, $\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)$ can be calculated in equation (18):

$$\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) - \left[1 - \left(\frac{\lambda^{API}}{\mu^{API}+\lambda^{API}}\right)^{z^{API}-1}\right] r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)$$

(18)

Simplifying, this produces Equation (19):

$$\tilde{r}^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{PL}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\lambda^{API}}{\mu^{API}+\lambda^{API}}\right)^{z^{API}-1}$$

(19)

That is, for raw material suppliers, the probability that a component failure leads to a system failure is the [probability that the plant-line system is available] multiplied by the [probability that all other suppliers are disrupted].

Manufacturing plants

Next, I will consider the susceptibility of the system to a particular manufacturing plant's failure, $\tilde{r}^{Plant}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)$. I will begin, similarly to the derivation for the API echelon, with the difference between the probability of the system working given a particular plant is available and the probability the system is working given a particular plant is unavailable (Equation (20)).

$$\tilde{r}_n^p\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right) = r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 1\right) - r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 0\right), \forall n \in N^p$$

(20)

To calculate the first term, I will consider the case whether a particular plant is available. The probability the system is available is the product of the probability at least one API supplier is available with the product that at least one plant-line combination is working (Equation (21)). The probability that at least one plant-line combination is available is one minus the probability

none are. This is the probability that none of the lines are available in the given plant multiplied by none of the other plant-line combinations are available.

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_n=1\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(1-\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right) \tag{21}$$

The probability the system is working given a particular plant is unavailable is given in Equation (22). It is the probability an API supplier is available multiplied by the probability at least one of the other plant-line combinations is available. The probability that at least one other plant-line combination is available is one minus the probability that no other plant-line combinations are available. That is, for the other z^p-1 plants, either the plant itself is unavailable (represented by the probability $\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)$), or all of the lines are unavailable in an available plant, which has a probability of $\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}$.

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_n=0\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(1-\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right) \tag{22}$$

Substituting these into equation (20) produces the following:

$$\tilde{r}^{Plant}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1} - \left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)$$

(23)

This simplifies to equation (24):

$$\tilde{r}^{Plant}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(1-\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)\left(\left(\frac{\lambda^p}{\mu^p+\lambda^p}\right)+\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}$$

(24)

Summarized, for manufacturing plants, the probability a component failure causes a system failure is the [probability that the supplier echelon is available] multiplied by the [probability the other plants are either not working or have no working lines] and multiplied by the [probability there is at least one working line in the given plant].

Manufacturing lines

Last, I will consider manufacturing lines. The probability the system fails given a particular line l fails is given by Equation (25).

$$\tilde{r}_n^l\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}\right) = r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 1\right) - r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}} \,\middle|\, \boldsymbol{z}, X_n = 0\right), \forall n \in N^l$$

(25)

If a given line is available, the system is available if a) there is a working supplier and b) either the plant corresponding to the line is available or the plant is unavailable and another plant-line combination is available. This probability is presented in Equation (26).

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_n=1\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\mu^p}{\mu^p+\lambda^p}+\frac{\lambda^p}{\mu^p+\lambda^p}\left(1-\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)\right) \quad (26)$$

If a given line is unavailable, the system is available if a supplier is available and a plant-line combination is available. The plant-line combination could be another line in the plant that corresponds to the given line or a different plant line combination. The probability is presented in Equation (27).

$$r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}|X_n=0\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\mu^p}{\mu^p+\lambda^p}\left(\left(1-\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}\right)+\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}\left(1-\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)\right)+\frac{\lambda^p}{\mu^p+\lambda^p}\left(1-\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)\right) \quad (27)$$

To calculate the probability that a given line's failure leads to a system failure, I substitute equations (26) and (27) into (25). This produces equation (28).

$$\tilde{r}^{Line}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)=r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}-\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}\left(1-\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)\right) \quad (28)$$

This can be simplified to Equation (29) and further to Equation (30).

$$\tilde{r}^{Line}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}\left(1-\left(1-\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1}\right)\right)\right) \quad (29)$$

$$\tilde{r}^{Line}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right) = r^{API}\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda}+\boldsymbol{\mu}}\right)\left(\frac{\mu^p}{\mu^p+\lambda^p}\right)\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l-1}\left(\frac{\lambda^p}{\mu^p+\lambda^p}+\frac{\mu^p}{\mu^p+\lambda^p}\left(\frac{\lambda^l}{\mu^l+\lambda^l}\right)^{z^l}\right)^{z^p-1} \quad (30)$$

The probability that a manufacturing line failure causes a system failure is then: the [probability that the supply echelon is available] multiplied by the [probability the plant the line is in is available] and multiplied by the [probability the other lines in the plant are unavailable and all other plants are unavailable].

3.4. Assumptions

The framework is subject to several assumptions. First, I assume that disruptions and the recovery processes occur independently at different components. The model also assumes that components within an echelon are homogenous; i.e., the same transition rates are applied. These follow the assumptions in earlier chapters (Chapters 2 and 3). Each of the API suppliers can fulfill the entire order of raw materials, and the manufacturing plants and lines can each produce sufficient quantities of the finished form to meet demand.

The model does not consider partial availability (cf. Yano and Lee (1995)) to follow what is observed in practice. For example, facilities may be closed because of natural disasters or quality issues (Palmer 2016, Thomas and Kaplan 2017).

To focus on capacity risk, the model does not consider the effects of holding safety stock. This is consistent with the low levels of safety stock held in practice for generic injectable drugs

(GAO 2016). A company could choose to maintain additional stock to meet demand during periods of system unavailability.

Finally, the aim of this model is to evaluate the vulnerability of a particular company's supply chain, and hence, I do not consider competition. If a given company is not able to meet demand, another company could supply the drug instead. However, drug shortages often affect drugs without competition or companies with large market shares (Fox et al. 2014). The case example considers a drug sold by a single company, where the unavailability of the supply chain would indicate a market-wide shortage if there is no inventory.

4. Numerical study

To illustrate the use of the model, I consider supply chain characteristics of one generic oncology drug. I evaluate the reliability of five types of supply chain configurations (lean; one backup supplier; one backup plant; one backup line; and one backup supplier and one backup plant) in different conditions. In the first analysis (Section 4.2), I will study the baseline case where disruptions and recovery are based on available data. The second analysi (Section 4.3) will be the "high-quality" case where the mean time-to-disruption is doubled for each of the components. The third analysis (Section 4.4) is the "quick recovery" scenario, where the mean time-to-recover is halved. In each analysis, the shortage results in the text will be rounded to the nearest percentage, and the mean time to status change will be rounded to the nearest 0.1 year.

4.1. Data

The data on a component's mean time to fail and recover are presented in Table 9. The failure data is based on the time between drug approval from the FDA (FDA 2018a) and the start of shortages as reported by the University of Utah Drug Information Service (UUDIS) (UUDIS 2016). The recovery data is based on the shortage durations reported by UUDIS (2016).

Table 9. Component characteristics

	Mean time to…	
Echelon	**Fail** $\left(\frac{1}{\lambda^{base}}\right)$	**Recover** $\left(\frac{1}{\mu^{base}}\right)$
Supplier	17.3 years	1.2 years
Plant	28.2 years	0.8 year
Line	8.5 years	0.08 years
Source	(FDA 2018a, UUDIS 2016)	(UUDIS 2016)

4.2. Baseline results

There are three primary outcomes: expected shortage, mean time-to-failure and mean time-to-recovery. These represent the overall proportion of time the drug is unavailable, the average time between shortages, and the average length of a shortage.

First, I will consider how the configuration of the supply chain affects expected shortages. The model uses the data from Table 9, and the results are presented in Table 10. If a manufacturer selects a supply chain without redundancy, shortages will occur 10% of the time. With back-up components, the expected shortages decrease. If they choose to maintain a back-up supplier, the expected shortage drops by over half to 4%. A back-up line is less effective, shortages are close to baseline at 9%. Maintaining two plants leads to expected shortages of 7%.

Table 10. Supply chain configurations and corresponding effects on drug shortages

Configuration			**Shortage**	**Mean time to…**	
Suppliers (z^{API})	**Plants** (z^{p})	**Lines per plant** (z^{l})	s	**System Failure** ($\bar{U}$)	**System Recovery** ($\bar{D}$)
1	1	1	10%	4.7 years	0.5 years
2	1	1	4%	6.2 years	0.3 years
1	2	1	7%	14.6 years	1.0 years
1	1	2	9%	10.5 years	1.0 years
2	2	1	1%	56.0 years	0.3 years

Figure 25 presents the relationship between configurations and shortages visually. It is a full-factorial analysis of maintaining between one and five suppliers, plants, and lines per plant. The color gradient of the points indicate the expected shortage from dark (higher shortages) to light (low shortages). As the number of components increase, the expected shortages decrease. As observed in Table 10, adding a second supplier decreases shortages substantially, where adding a second line does not have as large of an impact. As the number of components increases further, shortages continue to drop. Qualitatively, we can see that maintaining back-ups at multiple echelons is more effective at reducing shortages than having many back-ups at a single echelon.

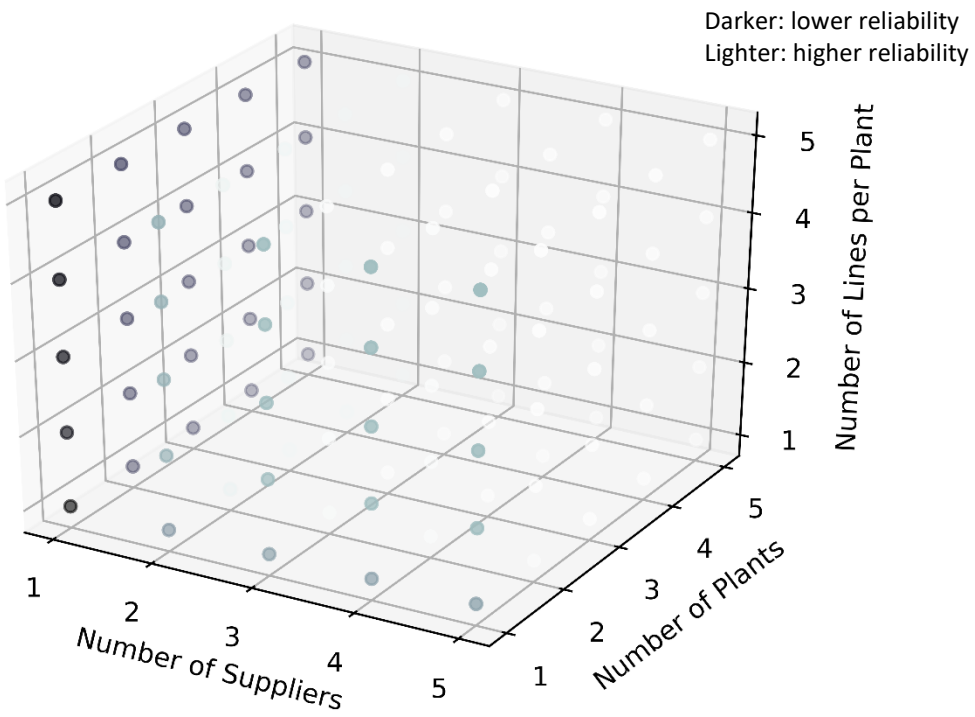

Figure 25. Configuration vs. expected shortages

We can also consider subsets of the data presented in Figure 25. By holding the number of components at one of the echelons constant, we can more precisely observe the effects of changing the number of components in the other echelons. These are presented in Figure 26 (single supplier), Figure 27 (single plant), and Figure 28 (single line).

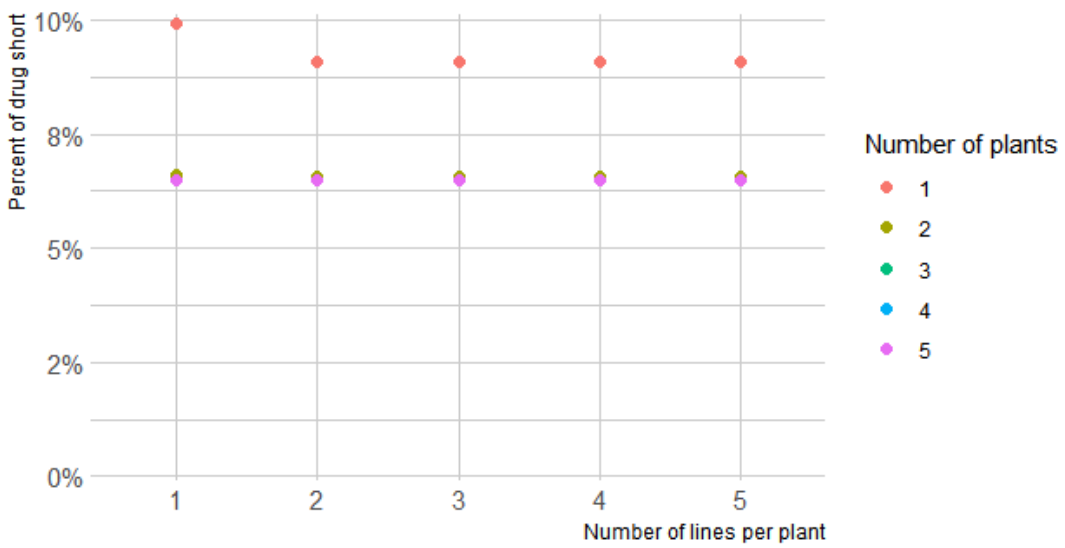

Figure 26. Effects of different configurations on shortages (single supplier)

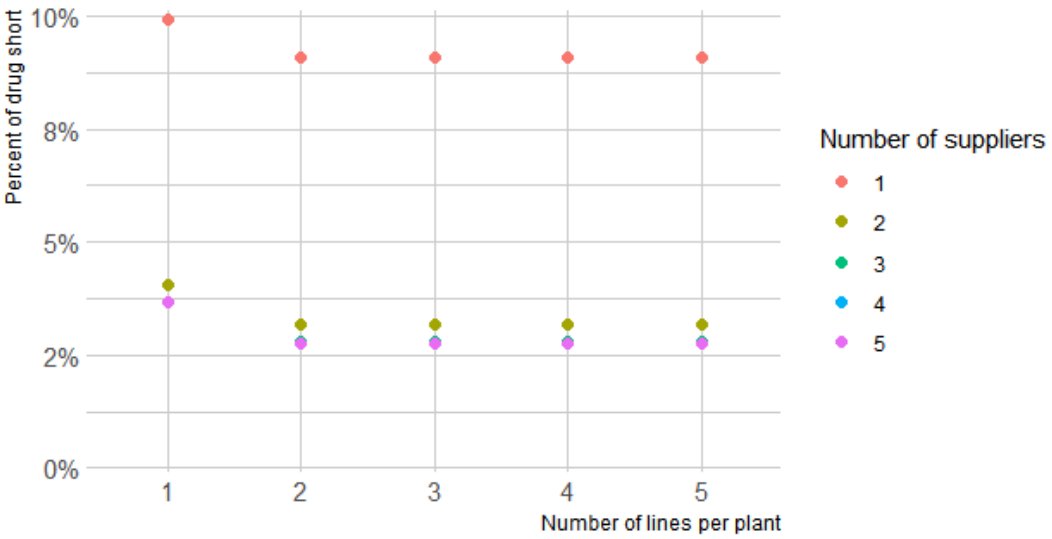

Figure 27. Effects of different configurations on shortages (single plant)

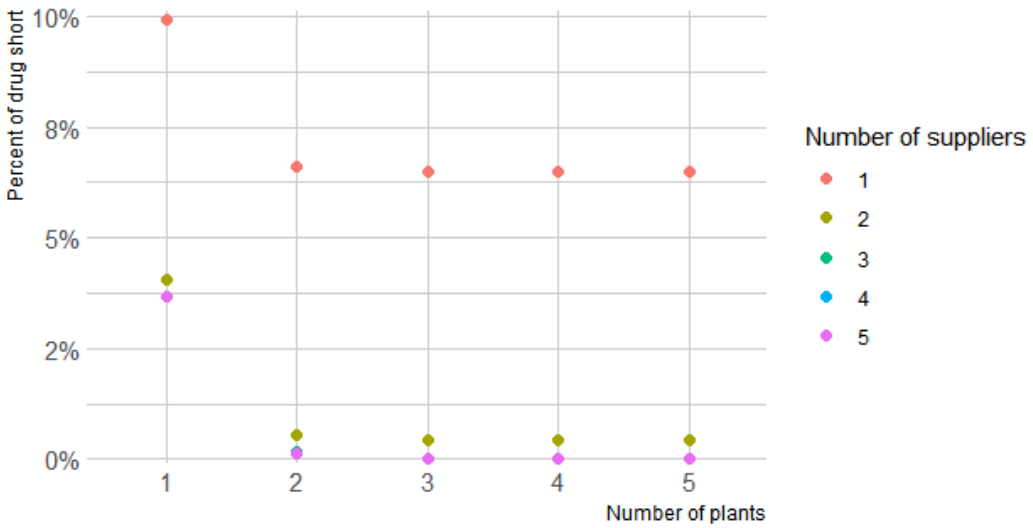

Figure 28. Effects of different configurations on shortages (single line)

In addition to expected shortages, the configurations also affect the times to system failure and system recovery. Recall that system unavailability is different from component unavailability. If a supply chain has two lines, both would need to be disrupted for the system to be unavailable.

In the case with a single component at each echelon, the mean time to failure is 4.7 years. This means that the average time between shortages is 4.7 years. As back-up components are added, the mean time-to-failure increases. The longest mean time-to-failure occurs when there is a back-up supplier and back-up plant (56 years). A back-up plant (with an implicit backup line) leads to a mean time-to-failure of 14.6 years. The system is ergodic, and with non-rounded values, the condition $\frac{\bar{U}}{\bar{U}+\bar{D}} = 1 - s$ holds. The model was also validated with a simulation model.

Note that configurations do not necessarily have the same effect on the expected shortage and the mean time-to-failure. A back-up supplier drops shortages by more than half (10% to 4%) and increases mean time to failure by 1.5 years (4.7 to 6.2 years). In contrast, a back-up line results in a small decrease in expected shortages vs. the no-back-up case (10% to 9%), but it

increases the mean time to failure by 5.8 years (4.7 to 10.5 years). In the back-up line case, shortages happen less often than in the back-up supplier case, but they last for 0.7 years longer when they do occur (1.0 vs. 0.3 years). This leads to a higher expected shortage for the back-up line case (9%) vs. 4% for suppliers even though shortages happen less frequently.

These results are caused by the differences in the mean time-to-recovery. The back-up line case has one of the longest mean times-to-recovery (1.0 years). When the system does fail, it is more likely to be caused by a supplier or plant failure than in the lean case. Suppliers and plants have a longer time to recovery (1.2 and 0.8 years, respectively) than lines (0.08 years). When they are more likely to be the cause of a shortage, the mean time-to-recovery for the supply chain overall will increase.

4.3. High-quality scenario

One proposal to reduce shortages is to increase the quality of the production process (FDA Drug Shortages Task Force 2019). If this were the case, disruptions would occur less often. For example, facilities would be shut down less frequently for quality violations and batches would be contaminated less frequently. This analysis will consider the effects of increasing component quality.

To model higher quality, the mean time to disruption are doubled relative to the baseline analysis. This represents a halving of the disruption rate; I apply $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\boldsymbol{quality}} = \frac{1}{2}\boldsymbol{\lambda}^{\boldsymbol{base}}$. The recovery rate remains the same as in the baseline analysis.

The results are presented in Table 11. For each configuration, the expected shortage is less than the baseline case (Section 4.2). This is intuitive because disruptions that could cause the system to be unavailable occur less frequently. The mean times-to-recovery are about the same as the baseline case.

Table 11. Shortage results with high-quality components

Configuration			**Shortage**	**Mean time to…**	
Suppliers	**Plants**	**Lines per plant**		**System Failure**	**System Recovery**
1	1	1	5%	9.5 years	0.5 years
2	1	1	2%	12.8 years	0.3 years
1	2	1	3%	31.5 years	1.1 years
1	1	2	5%	21.2 years	1.0 years
2	2	1	0%	214.1 years	0.3 years

4.4. Quick recovery scenario

Another opportunity to reduce shortages could be to improve the recovery process. This could involve reducing the time it takes to recover from a disruption. In this analysis, we half the mean time it takes to recover for each of the components. This represents a doubling of the recovery rate; $\boldsymbol{\mu} = \boldsymbol{\mu}^{\boldsymbol{quickrecovery}} = 2\boldsymbol{\mu}$.

The results of the quick recovery analysis are presented in Table 12. For each of the configurations, the expected shortages drop by about half relative to the baseline case. Similarly, the mean times-to-system-recovery are about half of baseline. The mean times-to-system-failure are about the same as in the baseline analysis, except the configuration with a backup supplier and plant which is doubled.

Table 12. Quick recovery results

Configuration			**Shortage**	**Mean time to…**	
Suppliers	**Plants**	**Lines per plant**		**System Failure**	**System Recovery**
1	1	1	5%	4.7 years	0.3 years
2	1	1	2%	6.4 years	0.1 years
1	2	1	3%	15.8 years	0.6 years
1	1	2	5%	10.6 years	0.5 years
2	2	1	0%	107.0 years	0.2 years

5. Pricing analyses

A natural question is what configuration would be most profitable. In an earlier chapter, stochastic programming models were developed to optimize the design of a pharmaceutical supply chain that may become disrupted (Chapter 2). The approach maximized the expected profit under different incentive policies and observed the optimal configuration and resulting shortages. It produced rough thresholds for when the optimal configuration would change.

The framework in this chapter is not an optimization model and does not include inventory as a resiliency strategy (as the SCDD-I model in Chapter 2 does). However, it can be used descriptively to compare the profitability and shortages under different policies. I will calculate the pricing thresholds at which the most profitable configuration changes.

As a case example, the analysis study vincristine sulfate, a drug used to treat pediatric cancers (Vincristine Sulfate 2018). The data on costs, pricing, and demand was previously presented in Chapter 2. For ease of reference, these data are reported again in Table 13 along with the notation for the objective function.

The objective of the company is to maximize its expected annual profit. It is modeled with Equation (31). It is comprised of the revenue for selling the drug; the variable costs for the

raw materials and production; and the fixed costs of the selected configuration and the program fee.

$$Q = d[(1-s)(q - c^{raw} - c^{prod})] - (f^{c,API} + c^{g,API})z^{API} - (f^{c,Plant} + c^{g,Plant})z^{p} - f^{c,Line}z^{p}z^{l} - f^{g,Program} \quad (31)$$

Table 13. Notation and data

Notation	Parameter/Outcome Name	Value		Source
	Annual fixed costs	*Company*	*GDUFA fees*	
$f^{c,API}; f^{g,API}$	Supplier	$33,000	$1,169	(FDA 2018b, Rudge 2012) and assumptions
$f^{c,Plant}; f^{g,Plant}$	Plant	$65,000	$4,401	
$f^{c,Line}$	Line	$32,500	n/a	
$f^{g,Program}$	Program fee		$9,700	(FDA 2018b, c)
c^{raw}	Raw material cost per ml	$0.34		(PharmaCompass 2018), procurement representatives
c^{prod}	Production cost per ml	$2.22		Calculated
q	Sales price per ml	$5.55		(IBM Micromedex 2018)
d	Annual demand in ml	90,000		(CMS 2018a, b, National Cancer Institute 2018)
Q	Expected annual profit			

§Costs in 2018 US dollars
GDUFA = Generic Drug User Fee Amendments

It has been suggested that the prices of drugs vulnerable to shortage may be too low (Frakt 2016). To study the potential effects of price changes on shortages, I evaluated the most profitable configuration under different policies. In particular, I calculated the expected profit for four supply chain configurations for prices between $0 and $30 per unit of the drug using equation (31). Prices were tested in increments of $0.25.

The profits are presented in Figure 29, and the four configurations are as follows. The "Lean" configuration is a single component in each echelon. "Two lines" represents a single API supplier and single plant with two lines. "Two plants" and "two suppliers" are defined similarly. The series "All" represents the configuration with a backup at each echelon.

As the price increases, the expected profit increases monotonically for each configuration. This is intuitive; for a given configuration, the expected quantity of the drug remains the same, and the revenue increases. This leads to higher profits. Below certain thresholds, though, the expected profit is \$0. In these cases, the company does not expect to make enough money to cover its expenses and would choose to not produce the drug. The threshold varies depending on the configuration because the configurations have different costs.

The most profitable configuration changes based on the price. The expected profit for the most profitable configurations are presented in Figure 30. Note that the unit prices considered in Figure 30 range from \$0 to \$50 (where the prices in Figure 29 are \$4 to \$10). As the price increases, it is optimal to choose a more reliable supply chain. At \$4.36, the company chooses to maintain a lean supply chain (one API supplier; one plant; and one line), and the expected shortage is 10%. At a unit price of \$9.06, it becomes more profitable to maintain a second supplier. The expected shortage is 4%. The next threshold is \$34.76 when the most profitable configuration is to have a backup at each echelon (two suppliers; two plants; one line in each plant). Expected shortages are 1%.

Equation (31) is used to calculate the profitability thresholds between configurations. The aim is to find the unit price, q, that produces the same expected profit. Below this value, the less reliable configuration is more profitable, and above this value, the more reliable configuration leads to higher profits.

I will take the example of the lean supply chain ($z^{API} = z^{p} = z^{l} = 1; s = 0.10$) vs. a supply chain with a backup supplier ($z^{API'} = 2; z^{p'} = z^{l'} = 1; s' = 0.04$). Substituting the configurations and expected shortages produces the equation (32):

$$d(1-s)(q - c^{raw} - c^{prod}) - 1(f^{c,API} + f^{g,API}) - 1(f^{c,Plant} + f^{g,Plant}) - 1f^{c,Line} - f^{g,Program}$$
$$= d(1-s')(q - c^{raw} - c^{prod}) - 2(f^{c,API} + f^{g,API}) - 1(f^{c,Plant} + f^{g,Plant}) - 1f^{c,Line} - f^{g,Program} \quad (32)$$

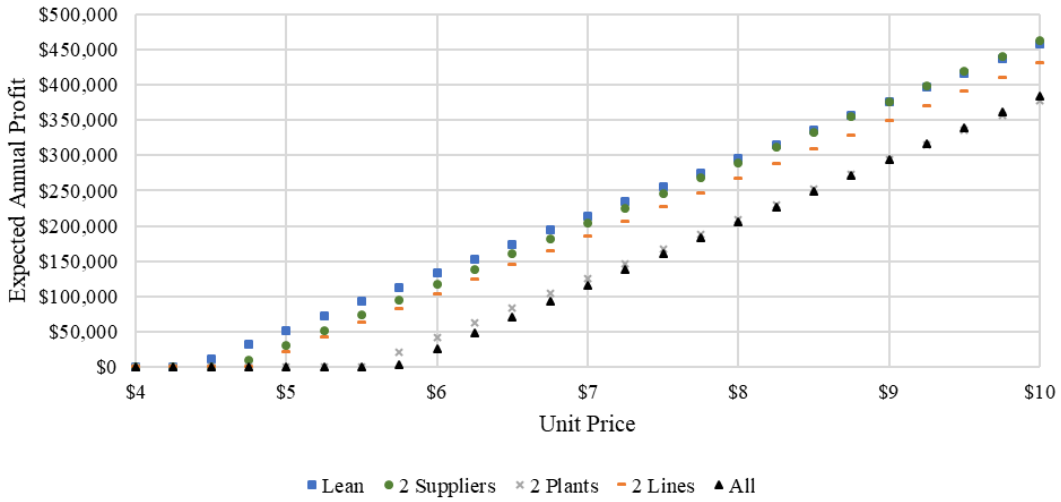

Figure 29. Profit of Configurations under Different Prices

Solving for the unit price q gives a value of $9.06. This is the threshold at which it becomes more profitable to maintain a backup supplier.

Drug companies are under increasing pressure to reduce prices (US Department of Health and Human Services 2018). At some point, it is no longer profitable for companies to produce the drugs.

Figure 30. Most Profitable Configuration by Price

To calculate the breakeven price, I set the profit in equation (31) to \$0 and solve for the breakeven unit price, q^0. The simplified expression is given in equation (33).

$$q^0 = c^{raw} + c^{prod} + \frac{(f^{c,API} + f^{g,API})z^{API} - (f^{c,Plant} + f^{g,Plant})z^p - f^{c,Line}z^p z^l - f^{g,Program}}{d * r\left(\frac{\boldsymbol{\mu}}{\boldsymbol{\lambda} + \boldsymbol{\mu}} \middle| \boldsymbol{z}\right)} \quad (33)$$

This represents the price that covers the variable costs ($u + v$) and the unit contribution to the fixed cost. The latter is the total fixed cost divided by the expected demand met.

For vincristine, the breakeven price for a lean supply chain is \$4.36; to have a backup supplier is \$4.64; and to have a backup supplier and plant is \$5.71. At each of these prices, the company would be covering its expected costs, and to be profitable, the company would need to charge more.

6. Discussion

There is a need to evaluate the resiliency of pharmaceutical supply chains to disruptions. I developed a closed-form model of supply chain availability and conducted analyses for generic oncology drugs.

The model requires minimal data and produces closed-form equations for system reliability, expected shortages, expected uptime, and expected downtime. It can be studied for an arbitrarily large number of components. The failure and recovery distributions can be represented by any continuous distribution; they are not limited to be exponential.

The two different types of supply chain stages (suppliers and plant-line combinations) provide a framework to extend the model to additional echelons. The reliability function is the product of the reliability of individual stages in series. To add another independent stage, the practitioner could multiply the reliability function by the reliability of the additional stage. If the additional stage were comprised of multiple echelons whose operation is dependent upon one another, the practitioner would calculate the reliability of the entire stage in a process similar to the plant-line combination.

A cancer drug, vincristine, was used to evaluate the reliability of sample supply chain configurations. Increasing the redundancy of the configuration decreases expected shortages. This varies by echelon; adding a supplier reduces shortages more than an additional line does (from a baseline of 10% to 4% and 9%, respectively). These support earlier results produced by a stochastic programming model (Chapter 2).

Adding redundancy also increases the expected time to a shortage. As the system has more capability to continue to produce during disruptions, shortages occur less often. The time to system failure is 4.7 years with the lean configuration and 6.2 years with an extra supplier.

There are mixed effects on shortage length, however. With a lean configuration, shortages persist for 0.5 years on average, and with an additional supplier, the expected length of the shortage is about half, 0.3 years. Yet with an additional line, shortages last for 1.0 years on average. This difference is because as redundancy is added, the echelon that is more likely to lead to a system failure changes. This analysis highlights the difference between time to recover and time to disruption.

The high-quality and quick-recovery analyses highlight other opportunities to reduce shortages. Many shortages are caused by quality issues (UUDIS 2016), and reducing the frequency of quality-related disruptions could reduce shortages. This could include upgrading equipment or taking steps to reduce contamination (ISPE 2015). It is worth considering ways to increase the time to disruption and decrease the time to recover. This may be particularly valuable if backup capacity is not viable, whether for cost or other reasons.

These results also underscore the importance of bringing production back quickly. Doubling the recovery rate could reduce shortages by half. This could occur through streamlining regulatory processes, often cited as burdensome (Tucker et al., 2020a).

The analytical model also allows us to conduct more precise policy analyses than are possible using a stochastic program. In particular, it can calculate the exact thresholds where the most profitable configuration changes. For example, a price threshold of $9.06 for vincristine may induce companies to add a backup supplier. The accuracy of the thresholds depends on the input data. Further sensitivity analysis would be necessary before it is applied in practice.

A company could use the model to evaluate the reliability of its own supply chain. The numerical study in this chapter is based on general estimates of the time to recover and disruption. If a company had more specific data on the characteristics of its specific

components, they could use the model to estimate more precisely the vulnerability of their own supply chains.

The company could also use the model to decide how much safety stock to hold to mitigate shortages. For example, they could carry inventory based on the expected shortage length. This would vary based on the configuration selected.

There are limitations to the use of the model. While only a handful of data points are needed, disruptions and recovery can be difficult to parameterize. The values were based on available databases, but in applying this method, the practitioner should recognize that the results are only as good as the underlying data. Data on pharmaceutical costs are frequently proprietary, and further sensitivity analyses are needed before policy results are implemented. Finally, the model does not allow for correlations between the components.

7. Conclusions

Improving supply chain reliability can help reduce drug shortages. My simple model provides metrics to evaluate configurations of pharmaceutical supply chains. It could be used by companies, regulators, or researchers to estimate shortages. The supply chain components have different costs, and there are opportunities to use the model to conduct cost-effectiveness analyses. Future work could consider extensions to include additional echelons or inventory.

CHAPTER V

Conclusions

Drug shortages are a crisis in the United States (US). New disruption-focused mathematical models and quantitative policy analysis were needed to address the problem. In this dissertation, I developed four mathematical models to tackle aspects of the crisis. Each considered different conditions under which decisions are made.

Summary of technical chapters

In Chapter 2, I compared policies proposed to reduce drug shortages. I presented two, new static supply chain design problems. They are among the first to include both disruptions and recovery within a design optimization. The formulation of the multi-stage stochastic program (that includes both configuration design and inventory decisions) implies the non-anticipativity constraints are redundant and simplifies the model considerably.

The model and analysis changed the perspective under which disruption problems are typically studied; the focus was on changing the underlying conditions to *incentivize* companies to be resilient. I observed that under status quo conditions for two example generic oncology drugs, it may be in a profit-maximizing company's best interest to select a supply chain configuration that is vulnerable to shortages. Yet, incentives could be put in place to change the optimal configuration; policies such as requiring back-up components in combination with moderate price increases could substantially reduce drug shortages.

In Chapter 3, I proposed a dynamic supply chain design model. The availability of the components is stochastic each period, and the time between disruptions and recovery apply different distributions, following available data. While this characteristic would previously have been intractable for a multi-stage stochastic program with binary state variables, I introduced a reformulation to apply the disruptions and recovery using stage-wise independent uncertainty. This set-up lays the groundwork for more realistic multi-stage disruption optimizations models; it expands our modeling capability from Bernoulli distributions to incorporate different geometrically-distributed disruption parameters as well.

Using the dynamic model, I studied the effects of reducing the lead times to add new supply chain components if disruptions occur. I found that for a sample generic injectable oncology drug, a lead time of 3 months could reduce shortages vs. a one year lead time. Halving the mean time to recover may reduce shortages as well. I also studied the effects of disruptions and product discontinuations. Rather than assuming the company will continue to produce the drug if a disruption occurs, the model allowed the decision-maker to drop out of the market. At baseline, in 12% of the simulations the company left the market. As the lead time decreases, product discontinuations also decrease (3% discontinuations with a 3 month lead time).

In Chapter 4, I developed a new model of pharmaceutical supply chain reliability (SCR). Using a given supply chain configuration, it outputs the probability the drug will be available and shortage characteristics (i.e., how often they occur and how long they last, on average). The baseline analyses considered generic oncology drugs and studied the reliability of different supply chain configurations. I found that adding a back-up line could double the time to shortage (from 4.7 years with a lean supply chain to 10.5 years with a back-up line). Adding redundancy at multiple levels could lead to even more substantial results. A back-up Active

Pharmaceutical Ingredient (API) supplier and back-up manufacturing plant could lead to the expected time to shortage of 56 years (compared to 4.7 years with a lean supply chain).

I also used the reliability model to analyze how improving component quality and reducing the time to recover from a disruption could reduce shortages. Either reducing the disruption rate (i.e., improving quality) or doubling the recovery rate (i.e., improving recovery) would be expected to drop the shortages by half. Further pricing analyses presented break-even points for the profitability of different configurations as well as the pricing thresholds for which design is optimal.

Extensions

Drug shortages continue to occur in the US and abroad. This dissertation lays a foundation to develop new mathematical models to address the problem. Several areas are worthy of further study.

The work in this dissertation focused on a single pharmaceutical company as the decision-maker. This approach was appropriate for my work because the example drugs I considered are produced by a single company. Yet, many drugs affected by shortages are generic (GAO 2016), and they may be affected by competition. It would be useful to consider how competition and the threat of disruptions affects supply chain design decisions. These analyses could provide evidence for regulatory policies to either incentivize additional competition or put further patent protection in place.

To analyze policies, I varied the values of exogenous input parameters. Modelers could develop bilevel models in which the leader (e.g., regulator) optimizes the policy parameters and the follower (e.g., a company) optimizes the supply chain design in response to the imposed

policies. A leader-follower context has been applied in other areas of policy but not for pharmaceuticals and disruptions.

Researchers could develop predictive models to identify which drugs are vulnerable to imminent shortages. Models to predict how long shortages will last could also be useful to help health systems plan their shortage management. Often manufacturer-provided projections for the time until shortage resolution are underestimated (McLaughlin et al. 2014).

Future work could refine the models to study increasingly realistic scenarios. With additional data (e.g., in-house at a pharmaceutical company), the design models could be extended to optimize location decisions under disruption and recovery. They could also incorporate correlations between facility disruptions.

Companies often make supply chain decisions for portfolios of drugs, rather than individual drugs. New models that include supply chain decisions for multiple drugs could be used to study trade-offs between maintaining the production of medically-necessary drugs and using manufacturing capacity to produce higher margin products.

This dissertation has focused on US policy, yet the pharmaceutical industry is international. Further research is necessary to consider international aspects of drug shortages. This could include optimizing regulatory policy among countries; how to allocate the drug when shortages occur; and pricing decisions.

To conclude, this dissertation developed new models to consider disruptions and recovery in pharmaceutical supply chain design. I applied the models to analyze policies to reduce drug shortages. It is my hope that the contributions will not be solely theoretical but also lead to the reduction in drug shortages.

APPENDIX

Supplemental Information for Chapter 2

Section 1 presents detail on the data used for the parameter values, and Section 2 presents the sensitivity and scenario analysis results. Section 3 includes proofs of the lemmas and theorem presented in the text.

1. Parameters

1.1. Distribution of time to recover

There are three recovery time distributions: time to supplier recovery, time to plant recovery, time to line recovery. The specific data on recovery times were unavailable. As a proxy, I used the distributions of shortage durations and adjusted them to account for reporting delays (UUDIS 2016).

For the supplier recovery distribution, I fit the distribution of shortage durations for all resolved shortages 2001-2016 in which the reported cause was due to raw material issues. For the plant recovery distribution, I fit the distribution of sole-source injectable shortage durations for which the cause was a manufacturing-related issue. I did not have data for the time to line recovery and assumed it was equal to 0.1 of the time to plant recovery. I evaluated this in sensitivity analysis. For each of these distributions, I also factored in the time the product is partially available (ASHP 2018) and evaluated several types of distributions and determined that an exponential distribution fit best (Delignette-Muller and Dutang 2015, R Core Team 2018). Because I consider discrete time periods, I discretized the distributions to apply a geometric distribution.

1.2. Distribution of time to disruption

To estimate the distribution of the time to disruption for suppliers, plants, and lines, I fit distributions for the time of FDA approval for the generic drug application to the date the shortage began (FDA 2018a, UUDIS 2016).

For suppliers, I considered drugs that were short due to raw material issues, and for plants, I considered shortages where the direct cause was a manufacturing issue. I assumed the time to line disruption was 0.3 times the time to a disruption of a plant. Similarly to the time to recover, I evaluated different types of distributions and fit geometric distributions for each.

1.3. Demand

I estimated the annual demand in the United States for vinblastine and vincristine based on Medicare Part B data for individuals at least 65 years old and the demographic information of the population that the drugs are used to treat.

The amounts of vinblastine and vincristine charged to Medicare Part B were approximately 45,000 mg in 2015 and 2016 (CMS 2018a). These drugs are commonly used to treat certain cancers (vinblastine - Hodgkin's disease, testicular cancer, and AIDS-related Kaposi's sarcoma; vincristine - Acute Lymphocytic Leukemia, Acute Myeloid Leukemia, Hodgkin's Disease, and Non-Hodgkin lymphoma; (Drugs.com 2018). The average proportion of new cases that are in individuals at least 65 years old are 13% for vinblastine-treated cancers and 51% for vincristine-treated cancers (National Cancer Institute 2018). Then I produced a rough estimate of the total national demand and converted to liquid volume (ml) based on the strengths provided in the Red Book (IBM Micromedex 2018).

Note that this process assumes the proportion of drug usage for these conditions by age is consistent with the proportion of new cases under 65 and that individuals under 65 are treated at

the same rate as individuals on Medicare. I evaluate the sensitivity of the results to the demand estimates in scenario analyses (Appendix Section 2).

1.4. Costs

GDUFA program fees:

Companies pay an annual fee to the FDA based on the number of approved Abbreviated New Drug Applications (ANDAs) they hold. These represent the number of generic drugs the companies are able to market. I estimated the GDUFA program fee that is allocated to each drug by dividing the appropriate fee by the total number of ANDAs the company holds based on the National Drug Code Directory (FDA 2018c). The company which produces vinblastine has 139 ANDAs which corresponds to a per drug cost of $11,445. The company which produces vincristine holds 164 ANDAs. This corresponds to a per drug cost of $9,700.

Fixed costs:

I estimated the plant fixed cost as the amortized 30-year cost of a plant based on estimates of costs of fill-and-finish facilities, adjusted to 2018 dollars, and divide by 100 products (GAO 2014, Rudge 2012). I assumed the fixed line cost is half of this value, and I assumed the supplier fixed cost is one quarter of the total raw material order (calculated as the variable raw material cost multiplied by annual demand).

Production costs:

The loaded cost to produce an injectable drug is 20-60% of the sales price based on conversations with a pharmaceutical manufacturing executive and estimates in literature (Jia and Zhao 2017). For the base parameter values, I assumed the loaded cost to be 40% of the sales price. I calculated the total cost to produce the drug annual as the product of 40%, the sales price, and annual demand. To back out the unit production costs, I subtracted the GDUFA and

non-GDUFA fixed costs and raw material costs and divided the resulting value by annual demand.

2. Sensitivity and scenario analyses

The one-way sensitivity analysis results for each drug are presented in the tornado diagrams of Figure A-1. In the upper bound (UB) analyses, the parameters were increased by 20% vs. the baseline values, and in the lower bound (LB) analyses, they were decreased by 20%.

In the baseline analysis, the time horizon length is two years. When this is varied to one year, three years, and five years, for vinblastine, the expected annual profit is within 1% of the baseline value, and the optimal solution does not change. Similarly, for vincristine, for one-, three-, and five- year time horizons, the expected annual profit is within 2% of the baseline value, and the optimal solution does not change. The baseline period length is two months. When the period length is decreased from two months to one month, the expected profit for vinblastine does not change and increases 3% vs baseline for vincristine. When the period length is increased to three months, the expected profit decreases 3% vs. baseline for both drugs. Adjusting the period length does not change the optimal solution from baseline for either drug. As the annual demand varies, the optimal solution does not change between 0.7-1.6 times baseline demand for vinblastine and 0.6-2 times baseline for vincristine. As the production capacity of the lines varies to 4 and 8 times the per-period demand, the companies continue to hold no inventory in the optimal solution.

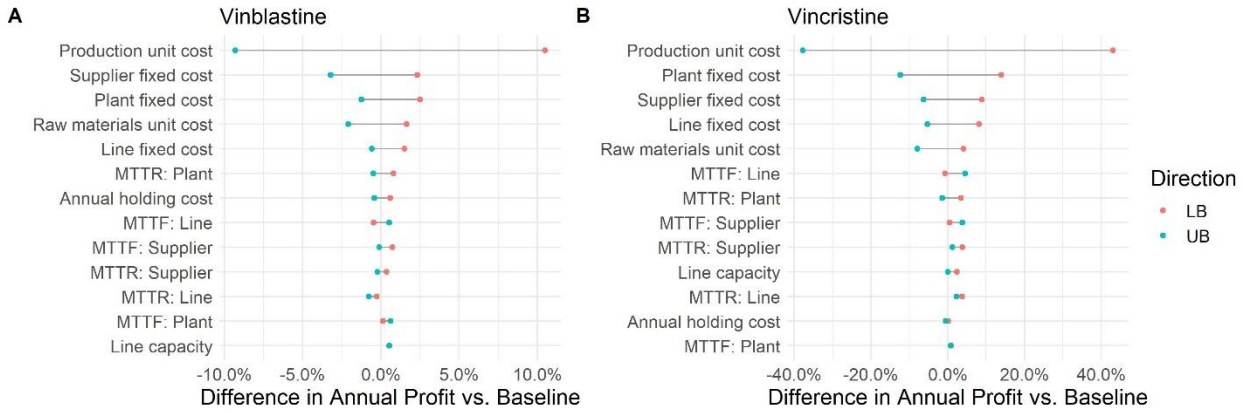

Figure 31. One-way sensitivity analysis results

3. Proofs

SCDD Model:

Lemma 1: $\theta_t^\omega \in \{0,1\}$ $\quad \forall t \in T, \omega \in \Omega$

$g^{Line} \in \mathbb{Z}^+$		By definition
$\xi_{nt}^\omega \in \{0,1\}$	$\forall n \in N, t \in T, \omega \in \Omega$	By definition
$\Rightarrow \theta_t^\omega \in \{0,1\}$		Constraints (6-11), Objective function (5)

■

SCDD-I Model:

Lemma 2: $C_t^\omega \in \{0,1\}$ $\quad \forall t \in T, \omega \in \Omega$ Constraints (22)

■

Lemma 3: $\tilde{C}_t^\omega \in \mathbb{Z}^+$ $\quad \forall t \in T, \omega \in \Omega$

$g^{Line} \in \mathbb{Z}^+$		By definition
$\xi_{nt}^\omega \in \{0,1\}$	$\forall n \in N, t \in T, \omega \in \Omega$	By definition
$\tilde{z}_{jl} \in \{0,1\}$	$\forall j \in J, l \in L$	Constraints (4e)
$\Rightarrow \tilde{C}_t^\omega \in \mathbb{Z}^+$ (23)	$\forall t \in T, \omega \in \Omega$	Lemma 2 and Constraints

■

Lemma 4: $I_0, I_t^\omega \in \mathbb{Z}^+$ $\quad \forall t \in \{0\} \cup T, \omega \in \Omega$

Constraints (10, 17, 20, 21, 25, 26, 27c), Objective function (15), Lemma 2

■

Lemma 5: $\theta_t^\omega \in \{0,1\}$ $\forall t \in T, \omega \in \Omega$ Constraints (10, 11, 16, 24), Objective function (15), Lemmas 2-4

■

Theorem 1: The following relationships are implied by SCDD-I.

$$C_t^\omega = C_t^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28a)$$
$$\tilde{C}_t^\omega = \tilde{C}_t^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28b)$$
$$I_t^\omega = I_t^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28c)$$
$$\theta_t^\omega = \theta_t^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28d)$$
$$\sum_{l \in L} v_{lt}^\omega = \sum_{l \in L} v_{lt}^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28e)$$
$$\sum_{j \in J} u_{jt}^\omega = \sum_{j \in J} u_{jt}^{\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28f)$$
$$\delta_t^{Avail,\omega} = \delta_t^{Avail,\omega'} \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28g)$$
$$\delta_t^{Suffic,\omega} = \delta_t^{Suffic,\omega'} \text{ except case: } \tilde{C}_t^\omega = I_0 - I_{t-1}^\omega \quad \forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega \quad (28h)$$

For (28e-f), I focus on the aggregate used in the objective function (15) because the specific supplier or line used in a given stage are not affected by uncertainty in later stages. For (28h), I exclude the case $\tilde{C}_t^\omega = I_0 - I_{t-1}^\omega$ because the indicator of sufficient capacity ($\delta_t^{Suffic,\omega}$) can be assigned either the value of 0 or 1 when the excess capacity ($\tilde{C}_t^\omega$) is equal to the safety stock deficit ($I_{t-1}^\omega - I_0$). The purpose of the indicator is to enforce the minimum operator in constraints (24). The values of the other variables in the constraints (24) are implied to be non-anticipative, and the indicator is not affected by uncertainty revealed in subsequent stages.

Proof of Theorem 1:

Lemma 6: $C_t^\omega = C_t^{\omega'}$ $\forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega$

If $\exists (j,k,l)$ s.t. $\xi_{jt}^\omega \xi_{kt}^\omega \xi_{lt}^\omega \tilde{z}_{jl} = 1$, then $C_t^\omega = 1$, else $C_t^\omega = 0$

$\forall j \in J, k \in K, l \in L_k, t \in T, \omega' \in S_t^\omega, \omega \in \Omega$ Constraints (22)

$\xi_{jt}^\omega = \xi_{jt}^{\omega'}; \xi_{kt}^\omega = \xi_{kt}^{\omega'}; \xi_{lt}^\omega = \xi_{lt}^{\omega'}$ $\forall j \in J, k \in K, l \in L, t \in T, \omega' \in S_t^\omega, \omega \in \Omega$ Definition of S_t^ω

$\Rightarrow C_t^\omega = C_t^{\omega'}$ $\forall \omega' \in S_t^\omega, t \in T, \omega \in \Omega$

■

Lemma 7: $\tilde{C}_t^{\omega} = \tilde{C}_t^{\omega'}$ $\quad \forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

Lemma 6 and Constraints (23)

■

Lemma 8: $\sum_{j \in J} u_{jt}^{\omega} = \sum_{l \in L} v_{lt}^{\omega}$ $\quad \forall t \in T, \omega \in \Omega$

$\sum_{l \in L} v_{lt}^{\omega} \leq \sum_{j \in J} u_{jt}^{\omega}$	$\forall t \in T, \omega \in \Omega$	Constraints (8)
$\Rightarrow \sum_{l \in L} v_{lt}^{\omega} = \sum_{j \in J} u_{jt}^{\omega}$	$\forall t \in T, \omega \in \Omega$	Objective function (15)

■

Lemma 9: If $C_t^{\omega} = 0$, then $\theta_t^{\omega} = \begin{cases} 1, & \delta_t^{Avail,\omega} = 1 \\ 0, & \delta_t^{Avail,\omega} = 0 \end{cases} \forall t \in T, \omega \in \Omega$

$\delta_t^{Avail,\omega} \in \{0,1\}$	$\forall t \in T, \omega \in \Omega$	Constraints (27c)
If $\delta_t^{Avail,\omega} = 1$, then $\theta_t^{\omega} = 1$	$\forall t \in T, \omega \in \Omega$	Constraints (10, 20)
If $\delta_t^{Avail,\omega} = 0$:		
$\xi_{jt}^{\omega} \xi_{kt}^{\omega} \xi_{lt}^{\omega} \tilde{z}_{jl} = 0$	$\forall t \in T, \omega \in \Omega, j \in J, k \in K, l \in L_k$	Constraints (22a)
$\sum_{l \in L} v_{lt}^{\omega} = 0$	$\forall t \in T, \omega \in \Omega$	Constraints (6-8, 14)
$\theta_t^{\omega} = 0$	$\forall t \in T, \omega \in \Omega$	Constraints (16, 24)

■

Lemma 10: If $I_{t-1}^{\omega} = I_{t-1}^{\omega'}$, then $\delta_t^{Avail,\omega} = \delta_t^{Avail,\omega'}$ $\quad \forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

$\nexists I_{t-1}^{\omega} \in (0,1)$	$\forall t \in T, \omega \in \Omega$	Lemma 4
$\delta_t^{Avail,\omega} = \begin{cases} 1, & I_{t-1}^{\omega} \geq 1 \\ 0, & I_{t-1}^{\omega} = 0 \end{cases}$	$\forall t \in T, \omega \in \Omega$	Constraints (25, 27c)

■

Lemma 11: If $\delta_t^{Avail,\omega} = \delta_t^{Avail,\omega'}$, then $\theta_t^{\omega} = \theta_t^{\omega'} \, \forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

$C_t^{\omega} \in \{0,1\}$	$\forall t \in T, \omega \in \Omega$	Lemma 2
Case 1: $C_t^{\omega} = 1 \Rightarrow \theta_t^{\omega} = 1$	$\forall t \in T, \omega \in \Omega$	Constraints (10, 19)
Case 2: $C_t^{\omega} = 0 \Rightarrow \theta_t^{\omega} = \theta_t^{\omega'}$	$\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$	Lemma 9

■

Lemma 12: $I_t^{\omega} = I_t^{\omega'}$ $\quad \forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

By induction:

$I_0^{\omega} = I_0^{\omega'}$	$\forall \omega' \in S_0^{\omega}, \omega \in \Omega$	Constraints (26)
$I_{t-1}^{\omega} = I_{t-1}^{\omega'}$	$\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$	Induction hypothesis

$\Rightarrow \sum_{l\in L} v_{lt}^{\omega} = \sum_{l\in L} v_{lt}^{\omega'}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Constraints (24) and Lemma 6 and 7

$\theta_t^{\omega} = \theta_t^{\omega'}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Lemmas 10 and 11

$\Rightarrow I_{t-1}^{\omega} + \sum_{l\in L} v_{lt}^{\omega} - \theta_t^{\omega} = I_{t-1}^{\omega'} + \sum_{l\in L} v_{lt}^{\omega'} - \theta_t^{\omega'}$
$\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

$\Rightarrow I_t^{\omega} = I_t^{\omega'}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Constraints (17)

■

Lemma 13: $\sum_{l\in L} v_{lt}^{\omega} = \sum_{l\in L} v_{lt}^{\omega'}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Constraints (24) and Lemmas 6, 7, and 12

■

Lemma 14: $\sum_{i\in I} u_{it}^{\omega} = \sum_{i\in I} u_{it}^{\omega'}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Lemmas 8 and 13

■

Lemma 15: $\delta_t^{Suffic,\omega} = \delta_t^{Suffic,\omega'}$ except case: $\tilde{C}_t^{\omega} = I_0 - I_{t-1}^{\omega}$
$\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$

$\delta_t^{\omega,Suffic} = \begin{cases} 1, & \tilde{C}_t^{\omega} > I_0 - I_{t-1}^{\omega} \\ 0, & \tilde{C}_t^{\omega} < I_0 - I_{t-1}^{\omega} \end{cases}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Constraints (10, 17, 18, 24, 27c)

$\Rightarrow \delta_t^{Suffic,\omega} = \delta_t^{Suffic,\omega'}$, except case: $\tilde{C}_t^{\omega} = I_0 - I_{t-1}^{\omega}$ $\forall \omega' \in S_t^{\omega}, t \in T, \omega \in \Omega$ Lemmas 6, 7, 12, and 13

■

Thus, Theorem 1 follows from Lemmas 6-15.

■

BIBLIOGRAPHY

BIBLIOGRAPHY

Adenso-Diaz B, Mena C, García-Carbajal S, Liechty M (2012) The impact of supply network characteristics on reliability. *Supply Chain Manag.* 17(3):263–276.

AHA, ASA, ASCO, ASHP, ASPEN, ISMP (2017) Letter to Congressman Griffith and Congresswoman DeGette.

Almansoori A, Shah N (2012) Design and operation of a stochastic hydrogen supply chain network under demand uncertainty. *Int. J. Hydrogen Energy* 37(5):3965–3977.

American Hospital Association (2011) AHA Survey on Drug Shortages. Retrieved (September 11, 2019), https://www.aha.org/system/files/content/11/drugshortagesurvey.pdf.

Anon (2012) FDASIA, Public Law 112–144. https://www.congress.gov/112/plaws/publ144/PLAW-112publ144.pdf.

Anon (2014) Health Policy Brief: Drug Shortages. *Health Aff.*

Anon (2018) Drugs.com. *[database online]*. Retrieved (July 6, 2018), https://www.drugs.com.

Anon (2018) Vincristine Sulfate. *Drugs.com*. Retrieved (July 6, 2018), https://www.drugs.com/monograph/vincristine-sulfate.html.

Anon (2018) Drug shortages roundtable: Minimizing the impact on patient care. *Am. J. Heal. Pharm.* 75(11):816–820.

Asbjørnslett BE (2009) Assessing the vulnerability of supply chains. *Int. Ser. Oper. Res. Manag. Sci.* 124:15–33.

ASHP (2013) *April 2013 Drug Shortages Summit Report: Evaluating Long-Term Solutions*

ASHP (2018) Drug Shortages FAQs. Retrieved (September 5, 2018), https://www.ashp.org/drug-shortages/current-shortages/drug-shortages-faqs.

Atan Z, Snyder L V. (2012) Inventory Strategies to Manage Supply Disruptions. Gurnani H, Mehrotra A, Ray S, eds. *Supply Chain Disruptions*. (Springer London, London), 115–139.

Azghandi R, Griffin J, Jalali MS (2018) Minimization of Drug Shortages in Pharmaceutical Supply Chains: A Simulation-Based Analysis of Drug Recall Patterns and Inventory Policies. *Complexity*:1–14.

Berk E, Arreola-Risa A (1994) Note on "future supply uncertainty in EOQ models." *Nav. Res. Logist.* 41:129–132.

Berry AJ (2014) Looking for the treatment for drug shortages: Not a simple prescription. *Mayo Clin. Proc.* 89(3):281–283.

Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: A Fresh Approach to Numerical Computing. *SIAM Rev.* 59(1):65–98.

Birge JR, Louveaux F (2011) *Introduction to Stochastic Programming* (Springer New York, New York, NY).

Bundschuh M, Klabjan D, Thurston DL (2003) Modeling robust and reliable supply chains. *Optim. Online e-print*:23.

Chabner BA (2011) Drug Shortages — A Critical Challenge for the Generic-Drug Market. *N. Engl. J. Med.* 365(23):2147–2149.

Chen SI, Fox ER, Hall MK, Ross JS, Bucholz EM, Krumholz HM, Venkatesh AK (2016) Despite federal legislation, shortages of drugs used in acute care settings remain persistent and prolonged. *Health Aff.* 35(5):798–804.

Cherici C, Frazier J, Feldman M, Gordon B, Petrykiw CA, Russell WL, Souza J (2011) *Navigating Drug Shortages in American Healthcare: A Premier healthcare alliance analysis* (Washington, D.C.).

Chopra S, Sodhi MS (2004) Managing risk to avoid supply-chain breakdown. *MIT Sloan Manag. Rev.* 46(1):53–61.

Church RL, Scaparra MP, Middleton RS (2004) Identifying critical infrastructure: The median and covering facility interdiction problems. *Ann. Assoc. Am. Geogr.* 94(3):491–502.

CMS (2018a) Medicare Enrollment Dashboard. Retrieved (May 7, 2018), https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMSProgramStatistics/Dashboard.html.

CMS (2018b) Part B National Summary Data File. Retrieved (May 7, 2018), https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Part-B-National-Summary-Data-File/Overview.html.

Conti RM (2011) An Economic Assessment of the Causes and Policy Implications of Current Specialty Drug Shortages. :1–10.

Craighead CW, Blackhurst J (2007) The Severity of Supply Chain Disruptions: Design Characteristics and Mitigation Capabilities. *Decis. Sci.* 38(1):131–156.

Dada M, Petruzzi NC, Schwarz LB (2007) A Newsvendor's Procurement Problem when Suppliers Are Unreliable. *Manuf. Serv. Oper. Manag.* 9(1):9–32.

Deleris L a., Erhun F (2005) Risk management in supply networks using Monte-Carlo simulation. *Proc. Winter Simul. Conf. 2005.*:1643–1649.

Delignette-Muller ML, Dutang C (2015) fitdistrplus: An R Package for Fitting Distributions. *J. Stat. Softw.* 64(4):1–34.

Dill S, Ahn J (2014) Drug shortages in developed countries - reasons, therapeutic consequences, and handling. *Eur. J. Clin. Pharmacol.* 70(12):1405–1412.

Doroudi R, Azghandi R, Feric Z, Mohaddesi O, Sun Y, Griffin J, Ergun O, et al. (2018) An Integrated Simulation Framework for Examining Resiliency in Pharmaceutical Supply Chains Considering Human Behaviors. Rabe M, A. A. Juan, Mustafee N, Skoogh A, Jain S, Johansson B, eds. *Winter Simul. Conf.* (IEEE), 88–99.

Dowson O, Kapelevich L (2017) SDDP.jl: a Julia package for stochastic dual dynamic programming. *Optim. Online*:1–24.

Fattahi M, Govindan K, Keyvanshokooh E (2017) Responsive and resilient supply chain network design under operational and disruption risks with delivery lead-time sensitive customers. *Transp. Res. Part E Logist. Transp. Rev.* 101:176–200.

FDA (2013) *Strategic plan for preventing and mitigating drug shortages*

FDA (2018a) Drugs@FDA: FDA Approved Drug Products. Retrieved (September 5, 2018), https://www.accessdata.fda.gov/scripts/cder/daf/.

FDA (2018b) Generic Drug User Fee Amendments. Retrieved (April 12, 2018), https://www.fda.gov/ForIndustry/UserFees/GenericDrugUserFees/default.htm.

FDA (2018c) National Drug Code Directory. Retrieved (April 12, 2018), https://www.fda.gov/Drugs/InformationOnDrugs/ucm142438.htm.

FDA Drug Shortages Task Force (2019) *Drug Shortages: Root Causes and Potential Solutions*

Fourer R, Gay DM, Kernighan BW (2002) *AMPL: A Modeling Language for Mathematical Programming* (Thomson Brooks/Cole, Pacific Grove, CA).

Fox ER, Sweet B V, Jensen V (2014) Drug Shortages: A Complex Health Care Crisis. *Mayo Clin. Proc.* 89(3):361–373.

Fox ER, Tyler LS (2013) Call to action: finding solutions for the drug shortage crisis in the United States. *Clin. Pharmacol. Ther.* 93(2):145–7.

Frakt A (2016) Sometimes, Drugs Are Not Costly Enough. *New York Times* (May 31) https://www.nytimes.com/2016/05/31/upshot/drug-prices-too-high-sometimes-theyre-not-costly-enough.html.

GAO (2011) GAO-12-116: Drug Shortages: FDA's Ability to Respond Should Be Strengthened. :1–63.

GAO (2014) GAO-14-194: Drug shortages. Public Health Threat Continues, Despite Efforts to Help Ensure Product Availability.

GAO (2016) GAO-16-595: Drug Shortages: Certain Factors Are Strongly Associated with This Persistent Public Health Challenge.

Gatesman ML, Smith TJ (2011) The Shortage of Essential Chemotherapy Drugs in the United States. *N. Engl. J. Med.* 365(18):1653–1655.

Gehrett BK (2012) A prescription for drug shortages. *J. Am. Med. Assoc.* 307(2):153–154.

Goldsack JC, Reilly C, Bush C, McElligott S, Bristol MN, Motanya UN, Field R, et al. (2014) Impact of shortages of injectable oncology drugs on patient care. *Am. J. Heal. Pharm.* 71(7):571–578.

Gottlieb S (2018a) Statement by FDA Commissioner Scott Gottlieb, M.D., on formation of a new drug shortages task force and FDA's efforts to advance long-term solutions to prevent shortages. Retrieved (July 18, 2018), https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm613346.htm.

Gottlieb S (2018b) Update on recovery efforts in Puerto Rico, and continued efforts to mitigate IV saline and amino acid drug shortages. *FDA Statement*. Retrieved (April 17, 2019),

https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm591391.htm.

Govindan K, Fattahi M, Keyvanshokooh E (2017) Supply chain network design under uncertainty: A comprehensive review and future research directions. *Eur. J. Oper. Res.* 263(1):108–141.

Guillén G, Mele FD, Espuña A, Puigjaner L (2006) Addressing the Design of Chemical Supply Chains under Demand Uncertainty. *Ind. Eng. Chem. Res.* 45(22):7566–7581.

Gupta DK, Huang SM (2013) Drug Shortages in the United States: A Critical Evaluation of Root Causes and the Need for Action. *Clin. Pharmacol. Ther.* 93(2):133–135.

Gurler U, Parlar M (1997) An Inventory Problem with Two Randomly Available Suppliers. *Oper. Res.* 45(6):904–918.

Gurobi Optimization LLC (2019) Gurobi Optimizer Reference Manual.

Ha C, Jun HB, Ok C (2018) A mathematical definition and basic structures for supply chain reliability: A procurement capability perspective. *Comput. Ind. Eng.* 120(February):334–345.

Hagspiel S (2018) Reliability with interdependent suppliers. *Eur. J. Oper. Res.* 268(1):161–173.

Haninger K, Jessup A, Koehler K (2011) *Economic Analysis of the Causes of Drug Shortages*

Hantel A, Siegler M, Hlubocky F, Colgan K, Daugherty CK (2019) Prevalence and Severity of Rationing During Drug Shortages. *JAMA Intern. Med.* 179(5):710–711.

Ho W, Zheng T, Yildiz H, Talluri S (2015) Supply chain risk management: A literature review. *Int. J. Prod. Res.* 53(16):5031–5069.

Hopp WJ, Yin Z (2006) Protecting Supply Chain Networks Against Catastrophic Failures. *Work. Pap.*

Hosseini R, Alsheikh M, Greene N, Seoane-Vazquez E, Harris C, Fox E, Rodriguez-Monguio R

(2018) Factors Associated With the Incidence of Oncology Drugs Shortages in the United States (2001-2017). *Value Heal.* 21:S372.

Hu X, Gurnani H, Wang L (2013) Managing risk of supply disruptions: Incentives for capacity restoration. *Prod. Oper. Manag.* 22(1):137–150.

IBM (2017) IBM ILOG CPLEX Optimization Studio 12.7.

IBM Micromedex (2018) RED BOOK. *[database online]*. Retrieved (April 26, 2018), www.micromedexsolutions.com.

IMS Institute for Healthcare Informatics (2011) Drug shortages: a closer look at products, suppliers and volume volatility. (November):36.

Infanger G, Morton DP (1996) Cut Sharing for Multistage Stochastic Linear Programs with Interstage Depency. *Math. Program.* 75(241–256):241–256.

ISPE (2015) *Drug Shortage Assessment and Prevention Tool* (Bethesda, MD).

ISPE, Pew Charitable Trusts (2017) *Drug Shortages: An Exploration of the Relationship between U.S. Market Forces and Sterile Injectable Pharmaceutical Products*

Jacobson SH, Sewell EC, Proano RA (2006) An analysis of the pediatric vaccine supply shortage problem. *Health Care Manag. Sci.* 9(4):371–389.

Jarosławski S, Azaiez C, Korchagina D, Toumi M (2017) Quantifying the persisting orphan-drug shortage public health crisis in the United States. *J. Mark. Access Heal. Policy* 5(1).

Jia J, Zhao H (2017) Mitigating the U.S. Drug Shortages Through Pareto-Improving Contracts. *Prod. Oper. Manag.* 26(8):1463–1480.

Jia X, Cui L (2012) Reliability research of k-out-of-n: G supply chain system based on copula. *Commun. Stat. - Theory Methods* 41(21):4023–4033.

Julka N, Baines T, Tjahjono B, Lendermann P, Vitanov V (2007) A review of multi-factor

capacity expansion models for manufacturing plants: Searching for a holistic decision aid. *Int. J. Prod. Econ.* 106(2):607–621.

Kaakeh R, Sweet B V., Reilly C, Bush C, DeLoach S, Higgins B, Clark AM, Stevenson J (2011) Impact of drug shortages on U.S. health systems. *Am. J. Heal. Pharm.* 68(19):1811–1819.

Karimi H, Ekşioğlu SD, Khademi A (2018) Analyzing tax incentives for producing renewable energy by biomass cofiring. *IISE Trans.* 50(4):332–344.

Kim S hyun, Cohen MA, Netessine S, Veeraraghavan S (2010) Contracting for Infrequent Restoration and Recovery of Mission-Critical Systems. *Manage. Sci.* 56(9):1551–1567.

Kim S, Scott Morton F (2015) A Model of Generic Drug Shortages: Supply Disruptions, Demand Substitution, and Price Control. *Work. Pap.*:1–36.

Kleindorfer PR, Saad GH (2005) Managing Disruption Risks in Supply Chains. *Prod. Oper. Manag.* 14(1):53–68.

Kleywegt AJ, Shapiro A, Homem-de-Mello T (2002) The Sample Average Approximation Method for Stochastic Discrete Optimization. *SIAM J. Optim.* 12(2):479–502.

Kweder SL, Dill S (2013) Drug shortages: The cycle of quantity and quality. *Clin. Pharmacol. Ther.* 93(3):245–251.

Lakhdar K, Papageorgiou LG (2008) An iterative mixed integer optimisation approach for medium term planning of biopharmaceutical manufacture under uncertainty. *Chem. Eng. Res. Des.* 86(3):259–267.

Le P, Seoane-Vazquez E, Rodriguez-Monguio R, Fox ER, Szeinbach SL, Dunehew AR, Montagne M (2011) The prevalence of pharmaceutical shortages in the United States. *J. Generic Med.* 8(4):210–218.

Lei X, MacKenzie CA (2019) Assessing risk in different types of supply chains with a dynamic

fault tree. *Comput. Ind. Eng.* 137(September):106061.

Lin JT, Chen TL, Chu HC (2014) A stochastic dynamic programming approach for multi-site capacity planning in TFT-LCD manufacturing under demand uncertainty. *Int. J. Prod. Econ.* 148:21–36.

Link MP, Hagerty K, Kantarjian HM (2012) Chemotherapy Drug Shortages in the United States: Genesis and Potential Solutions. *J. Clin. Oncol.* 30(7):692–694.

Liu J, Liu F, Zhou H, Kong Y (2016) An Integrated Method of Supply Chains Vulnerability Assessment. *Sci. Program.* 2016.

Losada C, Scaparra MP, O'Hanley JR (2012) Optimizing system resilience: A facility protection model with recovery time. *Eur. J. Oper. Res.* 217(3):519–530.

Luss H (1982) Operations Research and Capacity Expansion Problems: A Survey. *Oper. Res.* 30(5):907–947.

MacKenzie CA, Barker K, Santos JR (2014) Modeling a severe supply chain disruption and post-disaster decision making with application to the Japanese earthquake and tsunami. *IIE Trans.* 46(12):1243–1260.

Mak HY, Shen ZJ (2012) Risk diversification and risk pooling in supply chain design. *IIE Trans.* 44(8):603–621.

Marques CM, Moniz S, de Sousa JP, Barbosa-Póvoa AP (2017) A simulation-optimization approach to integrate process design and planning decisions under technical and market uncertainties: A case from the chemical-pharmaceutical industry. *Comput. Chem. Eng.* 106:796–813.

Martínez-Costa C, Mas-Machuca M, Benedito E, Corominas A (2014) A review of mathematical programming models for strategic capacity planning in manufacturing. *Int. J. Prod. Econ.*

153:66–85.

McCarthy S (2020) Coronavirus: China's factory closures could cause global medicine shortages. *South China Morning Post*.

McDaniels T, Chang S, Cole D, Mikawoz J, Longstaff H (2008) Fostering resilience to extreme events within infrastructure systems: Characterizing decision contexts for mitigation and adaptation. *Glob. Environ. Chang.* 18(2):310–318.

McLaughlin M, Kotis D, Thomson K, Harrison M, Fennessy G, Postelnick M, Scheetz MH (2013) Effects on patient care caused by drug shortages: a survey. *J. Manag. care Pharm.* 19(9):783–788.

McLaughlin MM, Pentoney Z, Skoglund E, Scheetz MH (2014) Projections for antiinfective drug shortages and time to actual resolution. *Am. J. Heal. Pharm.* 71(23):2074–2078.

Metzger ML, Billett A, Link MP (2012) The impact of drug shortages on children with cancer--the example of mechlorethamine. *N. Engl. J. Med.* 367(26):2461–3.

Mousazadeh M, Torabi SA, Zahiri B (2015) A robust possibilistic programming approach for pharmaceutical supply chain network design. *Comput. Chem. Eng.* 82:115–128.

Nagurney A (2006) *Supply Chain Network Economics: Dynamics of Prices, Flows and Profits* (Edward Elgar Publishing, Cheltenham, UK).

Narayana SA, Pati RK, Vrat P (2014) Managerial research on the pharmaceutical supply chain - A critical review and some insights for future directions. *J. Purch. Supply Manag.* 20(1):18–40.

National Cancer Institute (2018) Surveillance, Epidemiology, and End Results (SEER) Program. *SEER*. Retrieved (July 6, 2018), https://seer.cancer.gov/.

Nickel S, Saldanha-da-Gama F, Ziegler HP (2012) A multi-stage stochastic supply network

design problem with financial decisions and risk management. *Omega* 40(5):511–524.

Nishiguchi T, Beaudet A (1998) The Toyota Group and the Aisin Fire. *Sloan Manage. Rev.* 40(1):49–59.

O'Hanley JR, Church RL (2011) Designing robust coverage networks to hedge against worst-case facility losses. *Eur. J. Oper. Res.* 209(1):23–36.

Palmer E (2013) Updated: J&J's Doxil shortage to last until at least end of 2014. *FiercePharma*. Retrieved (March 7, 2020), https://www.fiercepharma.com/m-a/updated-j-j-s-doxil-shortage-to-last-until-at-least-end-of-2014.

Palmer E (2016) Teva halts production at sterile injectables plant to address FDA concerns. *FiercePharma*. Retrieved (January 30, 2017), https://www.fiercepharma.com/pharma/teva-halts-production-at-sterile-injectables-plant-to-address-fda-concerns.

Parlar M, Berkin D (1991) Future supply uncertainty in EOQ models. *Nav. Res. Logist.* 38(1):107–121.

Parlar M, Perry D (1996) Inventory models of future supply uncertainty with single and multiple suppliers. *Nav. Res. Logist.* 43(2):191–210.

Peng P, Snyder L V., Lim A, Liu Z (2011) Reliable logistics networks design with facility disruptions. *Transp. Res. Part B Methodol.* 45(8):1190–1211.

Pereira MVF, Pinto LMVG (1991) Multi-stage stochastic optimization applied to energy planning. *Math. Program.* 52(1–3):359–375.

Pettit TJ, Fiksel J, Croxton KL (2010) Ensuring Supply Chain Resilience: Development of a Conceptual Framework. *J. Bus. Logist.* 31(1):1–21.

PharmaCompass (2018) PharmaCompass. *[database online]*. Retrieved (May 17, 2018), www.pharmacompass.com.

Philpott AB, De Matos VL (2012) Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *Eur. J. Oper. Res.* 218(2):470–483.

Pires Ribeiro J, Barbosa-Povoa A (2018) Supply Chain Resilience: Definitions and quantitative modelling approaches – A literature review. *Comput. Ind. Eng.* 115(May 2017):109–122.

Qi L, Shen ZJM, Snyder L V. (2009) A Continuous-Review Inventory Model with Disruptions at Both Supplier and Retailer. *Prod. Oper. Manag.* 18(5):516–532.

De Queiroz AR, Morton DP (2013) Sharing cuts under aggregated forecasts when decomposing multi-stage stochastic programs. *Oper. Res. Lett.* 41(3):311–316.

R Core Team (2018) R: A language and environment for statistical computing. https://www.r-project.org/.

Reed BN, Fox ER, Konig M, Jackevicius CA, Masoudi FA, Rabinstein AA, Page RL (2016) The impact of drug shortages on patients with cardiovascular disease: causes, consequences, and a call to action. *Am. Heart J.* 175:130–141.

Roberts R, Ruthazer R, Chi A, Grover A, Newman M, Bhat S, Benotti S, et al. (2012) Impact of a national propofol shortage on duration of mechanical ventilation at an academic medical center. *Crit. Care Med.* 40(2):406–411.

Ross SM (2014) *Introduction to Probability Models* 11th ed. (Academic Press, Oxford, UK).

Rudge S (2012) The Cost of Pharmaceutical Facilities. *RMC Pharma*. Retrieved (September 5, 2018), http://rmcpharmanews.blogspot.com/2012/12/the-cost-of-pharmaceutical-facilities.html.

Saedi S, Kundakcioglu OE, Henry AC (2016) Mitigating the impact of drug shortages for a healthcare facility: an inventory management approach. *Eur. J. Oper. Res.* 251(1):107–123.

Saghafian S, Van Oyen MP (2012) The value of flexible backup suppliers and disruption risk

information: newsvendor analysis with recourse. *IIE Trans.* 44(10):834–867.

Sakli L, Hennet JC, Mercantini JM (2014) *An analysis of risks and vulnerabilities in supply networks* (IFAC).

Samvedi A, Jain V, Chan FTS (2013) Quantifying risks in a supply chain through integration of fuzzy AHP and fuzzy TOPSIS. *Int. J. Prod. Res.* 51(8):2433–2442.

Schmitt AJ (2011) Strategies for customer service level protection under multi-echelon supply chain disruption risk. *Transp. Res. Part B Methodol.* 45(8):1266–1283.

Schmitt AJ, Snyder L V., Shen ZJM (2010) Inventory systems with stochastic demand and supply: Properties and approximations. *Eur. J. Oper. Res.* 206(2):313–328.

Schmitt AJ, Sun SA, Snyder L V., Shen ZJM (2015) Centralization versus decentralization: Risk pooling, risk diversification, and supply chain disruptions. *Omega* 52:201–212.

Schmitt AJ, Tomlin B (2012) Sourcing strategies to manage supply disruptions. *Supply Chain Disruptions Theory Pract. Manag. Risk* 9780857297:51–72.

SeaRates (2018) Logistics Explorer. Retrieved (April 27, 2018), https://www.searates.com/reference/portdistance/.

Shah N (2004) Pharmaceutical supply chains: Key issues and strategies for optimisation. *Comput. Chem. Eng.* 28(6–7):929–941.

Shapiro A (2011) Analysis of stochastic dual dynamic programming method. *Eur. J. Oper. Res.* 209(1):63–72.

Sherwin MD, Medal H, Lapp SA (2016) Proactive cost-effective identification and mitigation of supply delay risks in a low volume high value supply chain using fault-tree analysis. *Int. J. Prod. Econ.* 175:153–163.

Sherwin MD, Medal HR, MacKenzie CA, Brown KJ (2020) Identifying and mitigating supply

chain risks using fault tree optimization. *IISE Trans.* 52(2):236–254.

Shishebori D, Snyder L V., Jabalameli MS (2014) A Reliable Budget-Constrained FL/ND Problem with Unreliable Facilities. *Networks Spat. Econ.* 14(3–4):549–580.

Simchi-Levi D, Schmidt W, Wei Y, Zhang PY, Combs K, Ge Y, Gusikhin O, Sanders M, Zhang D (2015) Identifying Risks and Mitigating Disruptions in the Automotive Supply Chain. *Interfaces (Providence).* 45(5):375–390.

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

Snyder L V., Daskin MS (2005) Reliability Models for Facility Location: The Expected Failure Cost Case. *Transp. Sci.* 39(3):400–416.

Snyder L V., Scaparra MP, Daskin MS, Church RL (2006) Planning for Disruptions in Supply Chain Networks. Johnson M, Norman B, Secomandi N, eds. *Tutorials Oper. Res.* (INFORMS, Hanover, MD), 234–257.

Snyder L V., Shen ZJM (2006) Supply and Demand Uncertainty in Multi-Echelon Supply Chains. *Work. Pap.*:1–46.

Song JS, Zipkin P (2009) Inventories with Multiple Supply Sources and Networks of Queues with Overflow Bypasses. *Manage. Sci.* 55(3):362–372.

Song JS, Zipkin PH (1996) Inventory Control with Information About Supply Conditions. *Manage. Sci.* 42(10):1409–1419.

Soni U, Jain V, Kumar S (2014) Measuring supply chain resilience using a deterministic modeling approach. *Comput. Ind. Eng.* 74(1):11–25.

Svensson G (2000) A conceptual framework for the analysis of vulnerability in supply chains. *Int. J. Phys. Distrib. Logist. Manag.* 30(9):731–750.

Tang CS (2006) Perspectives in supply chain risk management. *Int. J. Prod. Econ.* 103(2):451–488.

Tang O, Musa SN (2011) Identifying risk issues and research advancements in supply chain risk management. *Int. J. Prod. Econ.* 133(1):25–34.

Tang SY, Gurnani H, Gupta D (2014) Managing disruptions in decentralized supply chains with endogenous supply process reliability. *Prod. Oper. Manag.* 23(7):1198–1211.

Thanh PN, Bostel N, Péton O (2008) A dynamic model for facility location in the design of complex supply chains. *Int. J. Prod. Econ.* 113(2):678–693.

Thomas K, Kaplan S (2017) Hurricane Damage in Puerto Rico Leads to Fears of Drug Shortages Nationwide. *New York Times* (October 4) https://www.nytimes.com/2017/10/04/health/puerto-rico-hurricane-maria-pharmaceutical-manufacturers.html.

Thomas MU (2002) Supply chain reliability for contingency operations. *Proc. Annu. Reliab. Maintainab. Symp.*:61–67.

Tomlin B (2006) On the Value of Mitigation and Contingency Strategies for Managing Supply Chain Disruption Risks. *Manage. Sci.* 52(5):639–657.

Tucker EL, Cao Y, Fox ER, Sweet B V (2020) The Drug Shortage Era: A Scoping Review of the Literature 2001-2019. *Work. Pap.*

Tucker EL, Daskin MS, Sweet B V, Hopp WJ (2020) Incentivizing resilient supply chain design to prevent drug shortages: policy analysis using two- and multi-stage stochastic programs. *IISE Trans.* 52(4):394–412.

US Department of Commerce (2018) ACE TOOL: Shipping. Retrieved (April 27, 2018), https://acetool.commerce.gov/cost-risk-topic/shipping.

US Department of Health and Human Services (2018) American Patients First - The Trump Administration Blueprint to Lower Drug Prices and Reduce Out-of-Pocket Costs. (May).

UUDIS (2016) *Fox Data Summary.xlsx*

Vail E, Gershengorn HB, Hua M, Walkey AJ, Rubenfeld G, Wunsch H (2017) Association Between US Norepinephrine Shortage and Mortality Among Patients With Septic Shock. *J. Am. Med. Assoc.* 317(14):1433–1442.

Vilko JPP, Hallikas JM (2012) Risk assessment in multimodal supply chains. *Int. J. Prod. Econ.* 140(2):586–595.

Vizient (2019) *Drug shortages and labor costs*

Wagner SM, Bode C (2006) An empirical investigation into supply chain vulnerability. *J. Purch. Supply Manag.* 12(6 SPEC. ISS.):301–312.

Wagner SM, Neshat N (2010) Assessing the vulnerability of supply chains using graph theory. *Int. J. Prod. Econ.* 126(1):121–129.

Wagner SM, Neshat N (2012) A comparison of supply chain vulnerability indices for different categories of firms. *Int. J. Prod. Res.* 50(11):2877–2891.

Wiggins BS, Nappi J, Fortier CR, Taber DJ (2014) Cardiovascular Drug Shortages: Predominant Etiologies, Clinical Implications, and Management Strategies. *Ann. Pharmacother.* 48(9):1177–1186.

Woodcock J, Wosinska M (2013) Economic and technological drivers of generic sterile injectable drug shortages. *Clin. Pharmacol. Ther.* 93(2):170–176.

Wu T, Blackhurst J, O'Grady P (2007) Methodology for supply chain disruption analysis. *Int. J. Prod. Res.* 45(7):1665–1682.

Yano CA, Lee HL (1995) Lot Sizing with Random Yields: A Review. *Oper. Res.* 43(2):311–

334.

Yu X, Ahmed S, Shen S (2018) On the Value of Multistage Stochastic Facility Location with (or without) Risk Aversion. *Work. Pap.*:1–34.

Zou J, Ahmed S, Sun XA (2019) Stochastic dual dynamic integer programming. *Math. Program.* 175(1–2):461–502.