10-19-2010

Analysis of Stochastic Disruptions to Support Design of Capacitated Engineered Networks

Andrés Fernando Uribe-Sánchez
University of South Florida

Analysis of Stochastic Disruptions to Support Design of Capacitated Engineered Networks

by

Andrés Fernando Uribe-Sánchez

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Alex Savachkin, Ph.D.
Tapas K. Das, Ph.D.
José Zayas-Castro, Ph.D.
Alex Volinsky, Ph.D.
Yuncheng You, Ph.D.

Date of Approval
October 19, 2010

Keywords: enterprise networks, lean, capacity disruptions,
countermeasure policies, pandemic influenza

DEDICATION

To my father, Fernando Uribe.

ACKNOWLEDGEMENTS

I would like to thank Dr. Alex Savachkin, my friend, advisor and mentor. Without his support, this doctoral dissertation and rewarding research experience could not have been possible. Not only because his research interests were the reason to start my doctoral studies, but because his wisdom, hard work, work ethic, and truthful friendship have been a source of inspiration.

I want to thank Dr. Tapas K. Das for his friendship, mentoring and encouragement during the past years. I am profoundly thankful to Dr. Jose Zayas-Castro, for everything he has done for me and because he always believed in my professional potential. Thanks are due to my closest friends and graduate classmates at the University of South Florida. I also would like to thank the other members of my committee, Dr. Alex Volinsky and Dr. Yuncheng You, for their contribution of time and intellectual energy in this endeavor.

This achievement could not be possible without my family: Gladicilla, Anita, Mauro, Mapis, Paty, Hugo, Natis, Andrea, Sofi, and Pipe. I want to thank them for their love, support and encouragement throughout my life. They provided the strength I needed to continue. Thanks to God and to my father, because they have been always next to me. Last and by no means the least, I have been lucky to have the support, encouragement, faith, and love from my best friend and rock, Dayna Lee. This success is also yours.

TABLE OF CONTENTS

ABSTRACT

This work is a compilation of four manuscripts, three of which are published and one is in the second round of review, all in refereed journals. All four manuscripts focus on analysis of stochastic disruptions to support design of capacitated engineered networks. The work is motivated by limited ability to mitigate elevated risk exposure of large-scale capacitated enterprise networks functioning in lean environments. Such inability to sustain enterprise capacity in the face of disruptions of various origins has been causing multi-billion enterprise forfeitures and hefty insurance premiums. At the same time, decision support methodologies for reliable design of dynamic capacitated networks have been largely unavailable.

This work is organized as follows. *Paper 1* presents a methodology to analyze capacitated healthcare supply chains using a framework of forward flow-matching networks with multiple points of delivery. Special emphasis is given to developing stochastic models for capturing capacity trajectories at the points of delivery. *Paper 2* focuses on assuring capacity availability for a critical vertex exposed to random stepwise capacity disruptions with exponentially distributed interarrival times and uniformly distributed magnitudes. We explore two countermeasure policies for a risk-neutral decision maker who seeks to maximize the long-run average reward. We present an extensive numerical analysis as well as a sensitivity study on the fluctuations of some system parameter values. *Paper 3* extends the capacity assurance analysis for critical vertices by considering stepwise partial system capacity loss accumulating over time. We examine implementation of a countermeasure policy, aimed at reducing the disruption rate, for a risk-neutral decision maker who seeks to maximize long-run average return. We explore how the policy of maintaining the optimal disruption rate is affected by a number of system parameters. Finally, *Paper 4* presents a dynamic predictive methodology for mitigation of cross-regional pandemic outbreaks which can be used to estimate workforce capacity loss for critical vertices due to such societal disasters.

INTRODUCTION

Lean manufacturing philosophy and associated business practices have been widely embraced and deployed by global enterprises. The design of capacitated engineered networks is driven by lean manufacturing philosophy and implementation of global outsourcing, reduction of inventories, consolidation of suppliers, with the main purpose of improving operational efficiency. An example is the US automotive industry, where some estimates assert that the shift to JIT scheduling has saved companies more than $1 billion a year in inventory costs alone.

However, while lean manufacturing has substantially boosted operational efficiency, such reductionism has also left these enterprises operating in an increasingly risk-encumbered environment. Capacity disruptions triggered by forces of nature, property and process related hazards, and human interventions, have shown to have a profound impact on the engineered network risk. The following examples demonstrate how increased risk exposure and resulting capacity imbalance cause multi-billion enterprise forfeitures and increasing insurance premiums.

In 1995, an earthquake hit the port town of Kobe, razed to the ground 100,000 buildings and shut down Japan's largest port for over two years. In 1999, an earthquake in Taiwan displaced power lines to the semiconductor fabrication facilities responsible for more than 50 percent of the worldwide supplies of certain computer components, and shaved 5 percent off earnings for major hardware manufacturers including Dell, Apple, Hewlett-Packard, IBM, and Compaq. In September 2002, longshoremen on the US West Coast were locked out in a labor strike for 11 days, forcing the shutdown of 29 ports. With more than $300 billion of dollars in goods shipped annually through these ports, the dispute caused between $11 and $22 billion in lost sales, spoiled perishables and underutilized capacity. That December, a political strike in Venezuela made transnational businesses including GM, BP, Ford, Goodyear and Procter & Gamble halt their manufacturing for the duration

of the conflict. The 2003 outbreak of SARS in China and Singapore forced Motorola to close several plants. More recently, the 2005 hurricane Katrina destroyed the infrastructure of the state of Louisiana. In the same year, a labor strike in Asiana Airlines resulted in over 90 percent of the cargo services canceled. Man-made disasters, from terrorist attacks to computer viruses, are also on the rise.

Nowadays, a seemingly minor disruption in a lean infrastructure can have the potential to initiate a cascading sequence of capacity losses, threatening human lives and devastating large sectors of the economy. As a result of the above events, according to a recent survey by A.M. Best Company, Inc. of 600 executives, 69 percent of chief financial officers, treasurers and risk managers at Global 1,000 companies in North America and Europe view property-related hazards–such as fires and explosions–and supply chain disruptions as the leading threats to top revenue sources.

Historically, enterprises have lacked appropriate decision support methodologies and computational tools suitable for addressing risk incurred through capacity disruptions. In academia, traditional research efforts on minimizing the cost of supply chain operations and the focus on leveraging economies of scale often yield results that overconcentrate resources. Such optimal solutions can be very sensitive to system perturbations, initiated by internal and external disruptions. The inability to recognize the hidden costs of such overconcentration heightens the risk of increased costs and capacity imbalance.

The traditional literature that explicitly model the impact of disruptions has so far been focusing primarily on a *local* level of issues including scheduling, ordering, inventory management, and lot sizing. These works have modeled local entities exposed to *operational* instabilities in (i) production rate and lead times, (ii) supply rate, including machine failures, (iii) prices of resources, and (iv) process quality and yield.

One of the most common types of disruption appearing in the literature is that of supply rate changes. Important efforts include (i) management of stochastic demand systems, where the product supply is disrupted for periods of random duration, (ii) classic economic order quantity (EOQ) problem with supply disruptions, (iii) order-quantity/reorder-point inventory models with two suppliers subject to independent disruptions to compute the

exact form of the average cost expression, and (iv) analytical models for computing the stationary distribution of the on-hand inventory in a continuous-review inventory system with compound Poisson demand, Erlang distributed lead time, and lost sales, where the supplier can assume one of the two "available" and "unavailable" states at any point in time according to a continuous-time Markov chain.

Other studies address both supply disruptions and random demand. Some examples are (i) dynamic models concerning optimal inventory policies in the presence of market disruptions, which are often characterized by events with uncertain arrival time, severity and duration, (ii) the continuous-review stochastic inventory problem with random demand and random lead-time where supply may be disrupted due to machine breakdowns, strikes or other randomly occurring events, and (iii) the inventory-control model which includes a detailed Markovian model of the resupply system. A number of efforts which address supply and demand changes have been developed in the field of oil stockpiling, as there has been grave concern over the oil supply from the Middle East.

Modeling production rate disruptions (machine failures) has been largely addressed by extending classical economic manufacturing quantity (EMQ) models. These efforts include (i) the EMQ model when the production process is subject to a random deterioration from an in-control state to an out-of-control state, (ii) models of defect-generating process in the semiconductor wafer probe process to determine an optimal lot size, which reduces the average processing time on a critical resource, (iii) the approximation of the EMQ model with Poisson machine breakdowns and low failure rate, and (iv) the study of an unreliable production system with constant demand and random breakdowns, with the focus on the effects of machine failure and repair on optimal lot-sizing decisions. Other studies derive some unique properties of their model compared to the classical EMQ model, under the assumption of exponentially distributed time between failures and instantaneous repair of the machine.

Much of the recent literature focuses on minimizing costs of supply chain operations, whereas only a small fraction of the efforts have been dedicated to modeling the impact of various disruptions, such as those affecting demand patterns, supplier and production lead

times, prices, imperfect process quality, process yield, and other factors. However, so far, only a scant subset of the literature has attempted to analyze supply chain-wide disruptions, with most of the efforts focusing on the facility location problems.

At this point, we can summarize that research efforts addressing the disruption of supply are still comparatively new and scant, and none of these attempts have addressed the issue of stochastic capacity disruptions in directed networks, in closed form.

The work presented in this dissertation is motivated by the limited ability to mitigate elevated risk exposure of large-scale capacitated enterprise networks functioning in lean environments. Such inability to sustain enterprise capacity in the face of disruptions of various origins has been causing multi-billion enterprise forfeitures and hefty insurance premiums. At the same time, decision support methodologies for reliable design of dynamic capacitated networks have been largely unavailable. This doctoral dissertation is a compilation of four manuscripts, three of which are published and one is in the second round of review, all in refereed journals. All four manuscripts focus on analysis of stochastic disruptions to support design of capacitated engineered networks. This work is organized as follows.

The first manuscript is the paper titled "Analysis of Healthcare Supply Chain Systems Exposed to Random Capacity Disruptions". A final version of this document, to appear in the special issue on "*Healthcare Systems Engineering*" in the *International Journal of Collaborative Enterprise* by Inderscience Publisher, is presented in the Appendix B.1. Inderscience Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in this Ph.D. dissertation is attached in Appendix A. In this paper, we present an attempt to contribute to development of mathematical tools for modeling and analysis of risk inherent in health care supply chains, such as pharmaceutical and medical equipment/device enterprises. Our underlying formulation leverages the analytical convenience of formalism of capacitated feed–forward flow–matching networks (FMNs) with multiple points of delivery (POD). Special emphasis is given to developing stochastic models for capturing capacity trajectories at the points of delivery.

The second manuscript is titled "Two Countermeasure Strategies to Mitigate Random Disruptions in Capacitated Systems". This paper is published in the *Journal of Systems*

Science and Systems Engineering, volume 19, number 2, pages 210-226, 2010, by Springer Publisher. A printed version is presented in the Appendix B.2. Springer Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in this Ph.D. dissertation is attached in Appendix A. In this document, we focus on assuring capacity availability for a critical vertex. We examine random stepwise capacity disruptions with exponentially distributed interarrival times and uniformly distributed magnitudes. We explore two countermeasure policies for a risk-neutral decision maker who seeks to maximize the long-run average reward. A one-phase policy considers implementation of countermeasures throughout the entirety of a disruption cycle. The results of this analysis form a basis for a two-phase model which implements countermeasures during only a fraction of a disruption cycle. We present an extensive numerical analysis as well as a sensitivity study on the fluctuations of some system parameter values.

The third manuscript is titled "An Optimal Countermeasure Policy to Mitigate Random Capacity Disruptions in a Production System". This paper is published in the *International Journal of Agile Systems and Management*, volume 3, number 1/2, pages 4-17-226, 2008, by Inderscience Publisher. A printed version is presented in the Appendix B.3. Inderscience Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in this Ph.D. dissertation is attached in Appendix A. This work extends the capacity assurance analysis for critical vertices. We investigate a manufacturing system exposed to unpredicted capacity disruptions with exponentially distributed interoccurrence times and uniformly distributed magnitudes of disruptions. Each disruption renders a stepwise partial system capacity loss accumulating over time until the remaining capacity reaches a certain level, upon which the system gradually restores the lost capacity to the target level. We examine implementation of a countermeasure policy, aimed at reducing the disruption rate, for a risk-neutral decision maker who seeks to maximize long-run average return. We explore how the policy of maintaining the optimal disruption rate is affected by a number of system parameters.

The final manuscript is titled "A Predictive Decision Aid Methodology for Dynamic Mitigation of Influenza Pandemics". This document is currently in the second round of review

in the special issue on "*Optimization in Disaster Relief*" in the *OR Spectrum* by Springer Publisher. A final version of this document is presented in the Appendix B.4. Springer Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in this Ph.D. dissertation is attached in Appendix A. In this work, we present a large-scale simulation-based optimization methodology for developing dynamic predictive mitigation strategies for a network of regional pandemic outbreaks. The methodology considers measures of morbidity, mortality, and social distancing, translated into the societal and economic costs of lost productivity and medical expenses. We present a sensitivity analysis for estimating the marginal impact of changes in the total budget availability and variability of some critical mitigation parameters. The methodology is intended to assist public health policy makers. This effort can be used to estimate workforce capacity loss for critical vertices due to such societal disasters.

CONCLUSIONS

Current large-scale capacitated enterprise networks can span multiple continents and subsume hundreds of suppliers and customers. In addition, as these enterprises have been adopting the philosophy of lean manufacturing, including slimming inventory buffers, global outsourcing, and consolidating supplier base, such reductionism has also left them operating in an increasingly risk-encumbered environment. One of the most profound risk factors is unpredicted disruptions in available capacity, as might be caused by a number of controllable and uncontrollable factors, including nature, process hazards, and human activity. At the same time, efficient predictive analytics and computational tools suitable for analyzing the impact of capacity disruptions on network risk have been largely unavailable.

In this doctoral dissertation, motivated by limited ability to mitigate elevated risk exposure of large-scale capacitated enterprise networks functioning in lean environments, we focus on analysis of stochastic disruptions to support design of capacitated engineered networks. In what follows, we summarize the main contribution of each manuscript.

In the first manuscript, we present an original analysis of health care supply chain systems exhibiting converging assembly and exposed to random capacity disruptions. The supply chain was modeled as a feed–forward flow–matching network with multiple points of delivery - a mathematical formalism particularly useful for understanding changes in the aggregate network capacity, as might be impacted by unpredicted perturbations. In our novel work, we captured the time-fixed probability law on the available effective network capacity, in the presence of capacity propagation delays, in closed form. We then construed two models of stochastic dynamics of the available effective capacity at the network level. These trajectories can be used to model a number of disruptive scenarios.

Our analysis will contribute to understanding the degree of risk exposure and vulnerability of global health care supply chains. Moreover, our study can be used to provide a substantive measure of the trade-off between a lean structure of the supply chain and its

robustness and agility. The stochastic models of capacity disruptions discussed in this paper present one of the initial attempts to characterize supply chain dynamics via a FMN-based formalism. In the future, these stochastic models of capacity dynamics will be generalized to feature different recovery modes and incorporate disposition of strategic inventory buffers or capacity back-ups. This stochastic analytics will then be combined with the dynamics of demand at the points of delivery and capacity expenditures, to develop future methodologies, based on the tenets of utility theory, to provide decision support for design of resilient health care supply chain systems.

The second manuscript focuses on assuring capacity availability for a critical vertex of a large-scale capacitated enterprise network. We presented one of initial attempts to fill the vacuum in the existing literature and to focus on development of active countermeasure policies for managing lean capacitated systems in the presence of random capacity disruptions. The vertex under consideration experienced stepwise partial capacity disruptions with exponentially distributed interarrival times and uniformly distributed magnitudes, followed by instantaneous recovery. Examples of such capacity dynamics include: (i) shortage of repair personnel and performance degradation caused by failing equipment with a full repair upon a complete failure, (ii) non-self-announcing stepwise system failures, and (iii) gradual equipment phaseout and modernization.

This work explores two different countermeasure policies for a risk-neutral decision maker, who seeks to maximize the long-run average reward. The initial model considered a one-phase policy, where countermeasures were implemented during the entirety of a disruption cycle. The results of this model served as a basis to analyze a two-phase strategy, where countermeasures were activated during only a fraction of a disruption cycle. For the latter model, we aimed to determine the optimal threshold when the countermeasures should be disengaged. This paper provides one of the initial attempts for providing closed form solutions for optimal countermeasure policies for mitigation of random disruptions in capacitated systems. We hope that our work will be further generalized to address similar questions for capacitated systems evolving under more complex capacity dynamics.

The third manuscript extends the analysis on assuring capacity availability for a critical vertex of a large-scale capacitated enterprise network. We examined a production system experiencing periodic capacity disruptions, each of which is followed by a random recovery delay and a constant linear rate of recovery. The system manager's objective was to implement a countermeasure strategy to alleviate the rate of disruptions with a decreasing convex cost function. We derived an optimal level of disruption rate that maximizes the long-run average reward. The results of comparative statics suggest that choosing to maintain a lower disruption rate is optimal, if the system profitability is high. We concluded that higher unit profits and maximum capacity levels increase the costs of disruptions and hence, must be balanced by appropriate countermeasure strategies. Since shorter expected recovery delay and faster linear recovery reduce the economic loss of disruptions, the optimal level of lambda is increasing in both parameters.

Finally, the fourth manuscript presents a dynamic predictive methodology for mitigation of cross-regional pandemic outbreaks which can be used to estimate workforce capacity loss for critical vertices due to such societal disasters. The decision-aid methodology presented in this paper incorporates varying virus epidemiology and region-specific population dynamics. The model supports development of mitigation strategies for an efficient, progressive allocation of a limited resource budget over a network of regional outbreaks. The model seeks to dynamically minimize the impact of ongoing outbreaks and the expected impact of potential outbreaks, spreading from the ongoing regions. The methodology considers measures of morbidity, mortality, and social distancing, translated into the societal and economic costs of lost productivity and medical expenses.

APPENDIX A
COPYRIGHT APPROVALS

APPENDIX A (Continued)

Andres F. Uribe-Sanchez

From: Jeanette Brooks [jrb@inderscience.com]
Sent: Thursday, September 09, 2010 5:13 AM
To: auribe@mail.usf.edu
Cc: Jeanette Brooks
Subject: RE:

Dear Andres,

Thank you for your email. Inderscience is happy to give you the permission requested in your email below, provided full acknowledgement of the original source of publication is made clear, and a statement included that Inderscience retains the copyright.

Kind regards and good luck.!

Jeanette

Jeanette R Brooks (Dr)
Publications Director
Email: jrb@inderscience.com

IMPORTANT NOTICE - CHANGE OF ADDRESS NOTIFICATION
Please note that our mailing address has been changed:

<New address>
Inderscience Enterprises Limited
World Trade Centre Building II
29 route de Pre-Bois
Case Postale **856**
CH-1215 Geneve 15
Switzerland

From: M Dorgham
Sent: 08 September 2010 17:40
To: Jeanette Brooks
Subject: FW:

From: Andres F. Uribe-Sanchez [mailto:auribe@mail.usf.edu]
Sent: 08 September 2010 17:11
To: M Dorgham; Jim Corlett
Subject:

Dear Editor-in-Chief

I am currently a doctoral candidate in the Department of Industrial & Management System Engineering at the University of South Florida (Tampa, FL, U.S.). I work under the supervision of Prof. Alex Savachkin, and expect to successfully complete my Ph.D. degree by Fall 2010.

1

APPENDIX A (Continued)

Significant contributions of my doctoral research have been presented in the following Inderscience publications:

1. Savachkin A., **Uribe A.**, 2010, "Analysis of health care supply chain systems exposed to random disruptions", forthcoming in I. Journal of Collaborative Enterprise, special issue on "Healthcare Systems Engineering".

2. Savachkin A., Bakir N., **Uribe A.**, 2008, "An optimal countermeasure policy to mitigate random capacity disruptions in a production system", I. Journal of Agile Systems and Management, 3(1/2), 4-17.

Having already been granted authorization from my co-authors, I request a formal authorization to insert these two papers into my Ph.D. dissertation. The papers are attached in this email. A printed copy of your response and of the Inderscience author agreement document will be attached to my dissertation.

Best regards,

--

ANDRES URIBE-SANCHEZ, M.S.
Doctoral Candidate
Instructor & Research Assistant
Industrial and Management Systems Engineering
University of South Florida
e-mail: auribe@mail.usf.edu

2

APPENDIX A (Continued)

Andres F. Uribe-Sanchez

From: Essenpreis, Alice, Springer DE [Alice.Essenpreis@springer.com]
Sent: Friday, September 17, 2010 5:29 AM
To: auribe@mail.usf.edu
Subject: WG: Question from a Springer author / journal contributor

Flag Status: Flagged

Dear Mr. Uribe-Sanchez,

Thank you for your e-mail.

With reference to your request (copy herewith) to re-use material on which Springer controls the copyright, our permission is granted free of charge, on the following condition:
* full credit (journal title, volume, year of publication, page, article title, name(s) of author(s), figure number(s), original copyright notice) is given to the publication in which the material was originally published by adding: With kind permission of Springer Science+Business Media.

Kind regards,
_

Alice Essenpreis
Springer
Rights and Permissions
_
Tiergartenstrasse 17 | 69121 Heidelberg GERMANY
FAX: +49 6221 487 8223
permissions.Heidelberg@springer.com
WWW.springer.com/rights
_

-----Ursprüngliche Nachricht-----
Von: Pillai, Preeti, Crest Im Auftrag von Golive, Springer DE
Gesendet: Mittwoch, 8. September 2010 18:26
An: Permissions Heidelberg, Springer DE
Betreff: FW: Question from a Springer author / journal contributor

-----Original Message-----
From: SpringerAlerts@springeronline.com [mailto:SpringerAlerts@springeronline.com]
Sent: Wednesday, September 08, 2010 9:54 PM
To: Golive, Springer DE
Subject: Question from a Springer author / journal contributor

INTERNAL NAME - For Authors - Contact Springer ORIGINAL URL -
http://www.springer.com/authors?SGWID=4-111-12-673301-0

Additional information

FORM OF ADDRESS - Herr

1

APPENDIX A (Continued)

ACADEMIC TITLE - Doctoral Candidate
FIRST NAME - Andres
MIDDLE NAME - F
LAST NAME - Uribe-Sanchez
EMAIL - auribe@mail.usf.edu

For existing Springer authors:

JOURNAL_Title_DOI_JournalTitle - Two countermeasure strategies to mitigate random disruptions in capacitated systems,Niyazi O. Bakir, Alex Savachkin, Andres Uribe-Sanchez, J Syst Sci Syst Eng (Jun 2010) 19(2): 210-226 Comments - To whom corresponds:

I am currently a doctoral candidate in the Department of Industrial & Management System Engineering at the University of South Florida (Tampa, FL, U.S.). I work under the supervision of Prof. Alex Savachkin, and expect to successfully complete my Ph.D. degree by Fall 2010.

Significant contributions of my doctoral research have been presented in the following Springer publication:

1. Bakir N., Savachkin A., Uribe A., 2010, "Two countermeasure strategies to mitigate random disruptions in capacitated systems", Journal of Systems Science and Systems Engineering, 19(2), 210-226.

Having already been granted authorization from my co-authors, I request a formal authorization to insert this published paper into my Ph.D. dissertation. A printed copy of your response and of the Springer author agreement document will be attached to my dissertation.

Best regards,

--

ANDRES URIBE-SANCHEZ, M.S.
Doctoral Candidate
Instructor & Research Assistant
Industrial and Management Systems Engineering University of South Florida
e-mail: auribe@mail.usf.edu

2

APPENDIX B:
PUBLICATION 1: ANALYSIS OF HEALTHCARE SUPPLY CHAIN SYSTEMS EXPOSED TO RANDOM CAPACITY DISRUPTIONS

In this appendix, we present the final version of the manuscript "Analysis of Healthcare Supply Chain Systems Exposed to Random Capacity Disruptions" to appear in the special issue on "Healthcare Systems Engineering" in the International Journal of Collaborative Enterprise by Inderscience Publisher. The co-author, Dr. Alex Savachkin, authorized to include this document in my dissertation. Inderscience Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in my Ph.D. dissertation is attached in Appendix A.

APPENDIX B (Continued)

Analysis of healthcare supply chain systems exposed to random capacity disruptions

Alex Savachkin* and Andrés Uribe-Sánchez

Department of Industrial and Management Systems Engineering,
University of South Florida,
Tampa, FL 33620, USA
E-mail: alexs@usf.edu
E-mail: auribe@mail.usf.edu
*Corresponding author

Abstract: Capacity disruptions caused by forces of nature, process-related hazards, and man-made malignant interventions have shown to be a profound risk factor for lean healthcare supply chain systems, such as pharmaceutical and medical equipment/devices enterprises. At the same time, decision support methodologies suitable for analysing healthcare supply chain risk incurred through capacity disruptions have been largely unavailable. In this paper, we examine capacitated healthcare supply chain systems using a framework of feed-forward flow-matching networks with multiple points of delivery. We analyse two models of stochastic dynamics of such systems in the presence of random disruptions. The proposed closed-form analytics will serve as a foundation for future mathematical methodologies for risk-based design of resilient healthcare supply chains.

Keywords: healthcare; supply chain; lean; capacity disruptions; feed-forward, flow-matching networks; FMNs; stochastic dynamics.

Reference to this paper should be made as follows: Savachkin, A. and Uribe-Sánchez, A. (xxxx) 'Analysis of healthcare supply chain systems exposed to random capacity disruptions', *Int. J. Collaborative Enterprise*, Vol. X, No. Y, pp.000–000.

Biographical notes: Alex Savachkin received his PhD in Industrial Engineering at Texas A&M University in 2005. He is an Assistant Professor of the Department of Industrial and Management Systems Engineering at the University of South Florida, Tampa. His research interests include analytical support of enterprise risk analysis, public health disaster mitigation and emergency planning, and healthcare engineering.

Andrés Uribe-Sánchez is a Doctoral student of the Department of Industrial and Management Systems Engineering at the University of South Florida, Tampa. His research interests include risk-based analysis of random capacity dynamics in capacitated enterprise networks and development of dynamic mitigation strategies for cross-regional pandemic outbreaks.

APPENDIX B (Continued)

1 Introduction

In as much as the per capita healthcare expenses in the USA is six times more than that of other developed countries, the US healthcare system generates over a half-trillion dollars of system waste, annually (Porter and Teisberg, 2004), and was ranked 37th in the world by the World Health Organization (Ertas et al., 2000; Institute of Medicine, 2001). As pointed out by Institute of Medicine (IOM) in its 'Envisioning the National Healthcare Quality Report' (Hurtado et al., 2001), this alarming situation is mainly attributed to the inefficiency, poor communication, redundancy, and misuse of system resources. The IOM urged to reengineer the existing healthcare delivery system using a quality framework which includes six major dimensions:

1 safety

2 effectiveness

3 patient centredness

4 timeliness

5 efficiency

6 even-handedness.

Of the above six quality dimensions, two of them are directly impacted by pharmaceutical and medical equipment/devices supply chain systems: *efficiency*, defined as 'avoiding waste, including waste of equipment, supplies, ideas, and energy', and *timeliness*, that refers to "obtaining needed care and minimising unnecessary delays in getting that care" (Hurtado et al., 2001). According to the 'Global Pharmaceutical Market Forecast to 2012' report, the global pharmaceutical industry will earn around a trillion dollars in revenues, and the USA will still the key world importer (RNCOS, 2008). In 2003, the USA import of pharmaceutical goods reached almost $32 billion, with almost 79% of that coming from Europe, 13% from Asia, and 8% from North America (Davidson et al., 2005). Therefore, the improvement in the design of the underlying pharmaceutical and medical equipment/devices supply chain systems is viewed as one of main priorities in reengineering the USA healthcare delivery system (Brennan, 1998).

Nowadays, a typical pharmaceutical supply chain is a complex system that subsumes:

1 primary manufacturing that produces small quantities of an active ingredient as a chemical

2 secondary manufacturing that transforms the active ingredient into a final product

3 market warehouses or distribution centres

4 wholesalers

5 hospitals or retailers (Papageorgiou et al., 2001; Shah, 2004).

Since secondary manufacturing facilities, which serve regional markets, are often located in geographical regions different than the primary ones, the transportation (usually by ship) of the active ingredients can take weeks (Shah, 2004).

The design of modern pharmaceutical and medical equipment/devices supply chains is driven by lean manufacturing philosophy and implementation of global outsourcing,

APPENDIX B (Continued)

reduction of inventories, and consolidation of suppliers, with the main purpose of improving operational efficiency. However, such reductionism has also left these enterprises operating in an increasingly risk-encumbered environment (Booth, 1996; Mohd, 2006). Capacity disruptions triggered by forces of nature, property and process related hazards, and human interventions, have shown to have a profound impact on the supply chain risk. The following examples demonstrate how increased risk exposure and resulting capacity imbalance cause multi-billion enterprise forfeitures and insurance premiums.

In 1995, an earthquake devastated Kobe, the largest Port of Japan, cutting the supply of pharmaceutics and medical equipment/devices shipped through this port for two years (Horwich, 2000; Sarkisian, 1995). Another example of such radical single-point disruption is the September 2002 labour strike on the US West Coast when 29 ports were locked out for 11 days (Hall, 2004; Isidore, 2002). More recently, the 2005 hurricane Katrina destroyed an estimated of 106 health centres in the affected states, impacting among others, the distribution of medicines (Rosenbaum, 2006). The American Diabetes Association estimated that over 205,000 people with diabetes were affected by the insulin shortage due to hurricanes Katrina and Rita (Coons, 2005). Same year, a labour strike in Asiana Airlines resulted in over 90% of the cargo services cancelled (M2 Comm., 2005). Man-made disasters are also on the rise, from terrorist attacks to computer viruses (Lemos, 2003a, 2003b). Nowadays, a seemingly minor disruption born upstream a lean healthcare supply chain can have the potential to initiate a cascading sequence of capacity losses, threatening human lives and devastating large sectors of the economy.

The existing literature on the design of pharmaceutical supply chains is scarce, disconnected, and mostly in the form of qualitative discussions. Description and analysis of the key issues can be found in Shah (2004), Booth (1996), Harland (1996), She (2001), and Koh et al. (2003). The integration of suppliers in the design of new pharmaceutical supply chains has been discussed in Petersen et al. (2005). Allocation of production plants and distributions points in the form of an optimisation problem is presented in Papageorgiou et al. (2001). The impact of dynamic demand on production and inventory policies is debated in Booth (1996). Finally, Bayer et al. (2007) discussed the impact of uncertainty about future requirements due to technological, demographic, medical and policy changes.

In some ways, pharmaceutical supply chains resemble the ones associated with consumer goods (Koh et al., 2003). In the traditional supply chain design literature, the focus on leveraging the economies of scale can yield results that over-concentrate resources. The optimal solutions can then be sensitive to system perturbations, initiated by internal and external disruptions. As a result, the inability to consider the hidden cost of over-consolidation heightens the risk of capacity imbalance. The traditional literature that explicitly models the impact of disruptions has so far been focusing primarily on a local level of issues including scheduling, ordering, inventory management, and lot sizing. These works have been mainly modelled local entities exposed to operational instabilities in:

1 production rate and lead times (Arreola-Risa and DeCroix, 1998; Meyer et al., 1979; Parlar and Berkin, 1991; Parlar and Perry, 1996; Posner and Berg, 1989)

2 supply rate (Bielecki and Kumar, 1988; Parlar, 1997; Song and Zipkin, 1996), including machine failures (Abboud, 1997; Buzacott and Shantikumar, 1993;

APPENDIX B (Continued)

Gallego, 1988a, 1988b; Groenevelt et al., 1992; Henig and Gerchak, 1990; Hopp et al., 1989; Kim and Hong, 1997; Lee, 1992; Porteus, 1986; Rosenblatt and Lee, 1986)

3 prices of resources (Arcelus and Srinivasan, 1995; Ardalan, 1995; Aull-Hyde, 1992; Taylor and Bradley, 1985; Tersine and Barman, 1995)

4 process quality and yield (Chao, 1987; Mohebbi, 2003; Weiss and Rosenthal, 1992).

However, so far, only a scant subset of the literature has attempted to analyse supply chain-wide disruptions, with most of the efforts focusing on the facility location problems. Notable examples include Eiselt et al. (1996), Snyder and Daskin (2005), Savachkin et al. (2008), Snyder et al. (2007), Lee (2001), Tsiakis et al. (2001), and Atamt and Zhang (2007). None of these attempts have addressed the issue of stochastic capacity disruptions in directed networks, in closed form.

In this paper, we present an attempt to contribute to development of mathematical tools for modelling and analysis of risk inherent in healthcare supply chains, such as pharmaceutical and medical equipment/devices enterprises. Our underlying formulation leverages the analytical convenience of formalism of capacitated feed-forward flow-matching networks (FMNs) with multiple points of delivery (POD). FMNs offer a mathematical structure suitable for modelling network assemblies, where, at each tier, the flows of system entities (e.g., materials, components, etc.) are combined and forwarded for subsequent merging. This formalism can be particularly useful for understanding dynamics in the capacitated healthcare supply chains impacted by unpredicted disruptions. In this effort, we are primarily focused on examining the capability of healthcare supply chains to sustain anticipated demand for system throughput. Our results are useful for understanding, up to the probability law, the impact of unpredicted capacity disruptions in healthcare supply chains. Special emphasis is given to developing stochastic models for capturing capacity trajectories at the POD which will provide the analytical basis necessary to model network risk and develop risk-based design methodologies.

The paper is organised as follows. Section 2 introduces the terminology and notation, provides a necessary account of the feed-forward FMN formalism, and constructs the probability law on the available effective capacity for FMNs with multiple POD, in the presence of capacity propagation delays. Section 3 builds upon this analysis and presents two stochastic models of the available effective network capacity. First, in Section 3.1, we discuss an instantaneous capacity loss with random recovery delay and a constant recovery rate trajectory. A special case of this model is discussed in Section 3.2. Finally, the concluding remarks are presented in Section 4.

2 Feed-forward FMNs with multiple POD

Feed-forward FMNs offer a convenient mathematical paradigm that can describe dynamics of a wide range of capacitated networks which exhibit a converging assembly flow. The topology of such tier-like systems is such that, at each tier, system entities are joined in some way and advanced to subsequent merging. Supply chain systems of global pharmaceutical and medical equipment/devices enterprises present a notable example of such systems, where each vertex (i.e., a production facility) combines the input of materials, components, and parts from the suppliers in the preceding tier, to produce more

APPENDIX B (Continued)

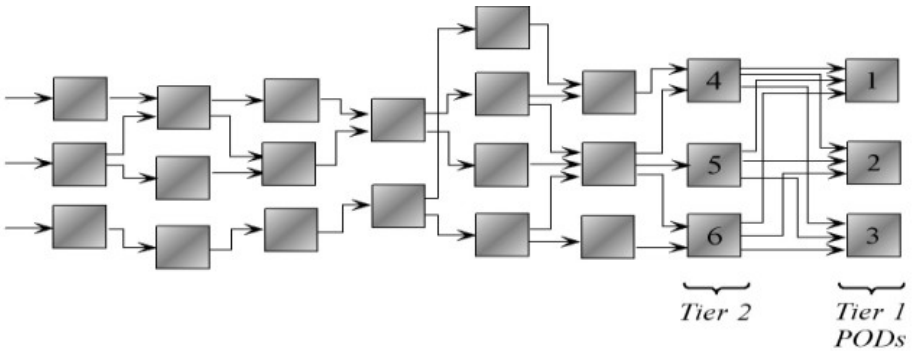
complex sub-assemblies. These sub-assemblies are then advanced to the succeeding tier, where further merging occurs. The converging assembly flow nature prescribes that the in-flow rates for any vertex must match. A final product emerges the network at the POD. At a strategic level of network design, re-entrant flows may be assumed as insignificant and hence, the flow of assembly can be assumed as unidirectional (or feed-forward). The reader is referred to Dorogovtsev and Mendes (2003) for a more detailed exposition of feed-forward networks.

In what follows, we introduce the terminology and notation, followed by derivation of the time-fixed distribution of the available effective capacity of FMN with multiple POD.

2.1 Terminology and notation

Consider a feed-forward network, such as one shown in Figure 1, which has a finite number $n \geq 2$ of vertices and $m \geq 2 (m \leq n)$ non-empty tiers.

Figure 1 A feed-forward network with multiple POD



The tiers are numbered in ascending order, starting from tier 1, where $D > 1$ POD are located, and moving upstream the network. Vertices are numbered in ascending order, starting from vertices in tier 1, from top to bottom within a tier, and moving upstream the network. Let N_k be the set of vertices that belong to tier $k, k = 1, ..., m.$ Let $N = \bigcup_{k=1}^{m} N_k.$

In what follows, we assume that:

1 each unique component (ingredient, part, material, etc.) is supplied by only one vertex

2 each vertex performs a single merging operation. We introduce the following terminology and notation.

Throughput: the long-run average number of units of finished product per unit time. For ease of exposition, the throughput of all vertices is measured in terms of the equivalent units of the finished product.

Available production capacity of vertex j at time $t > 0, C_{p_j}(t)$: maximum throughput that production resources of vertex j are capable of sustaining at t.

Available supply capacity of vertex j at time $t > 0, C_{s_j}(t)$: maximum throughput that the supply to vertex j is capable of sustaining at t.

APPENDIX B (Continued)

Available effective capacity of vertex j *at time* $t>0, C_{e_j}(t)$: maximum throughput of vertex j at t. Note that $C_{e_j}(t)=\min\{C_{p_j}(t), C_{s_j}(t)\}$.

Multifurcation coefficient, $0 \le A_{ji} \le 1, j>i$: the proportion of the available effective capacity of vertex j designated to feed vertex $i (\sum_i A_{ji}=1)$.

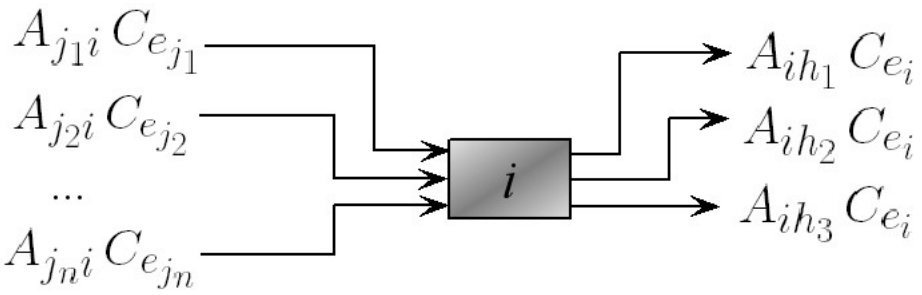
Capacity propagation delay, $T_{ji} \ge 0$: time in-transit between vertices j and i.

We say that a feed-forward network is *flow-matching*, if for any $t>0, i \in N_k, j_n \in N_{k+1}, k=1,...,m,$ the following balance equation holds (see Figure 2):

$$A_{j_1 i} C_{e_{j_1}}(t-T_{j_1 i}) = A_{j_2 i} C_{e_{j_2}}(t-T_{j_2 i}) = ... = A_{j_n i} C_{e_{j_n}}(t-T_{j_n i}) = \sum_{h_r \in N_{k-1}} A_{i h_r} C_{e_i}(t). \quad (1)$$

The left-hand side of equation (1) states the flow-matching assembly principle, i.e., it asserts that for any network vertex, the effective in-flow capacities from the feeding suppliers must match or else, they create a capacity imbalance; the right hand side of the equation essentially states the capacity preservation principle: each effective in-flow must be equal to the total effective out-flow. The flow-matching nature of assembly can be impacted by the presence of unpredicted capacity disruptions of various origins. Analysis of this impact will be the focus of the forthcoming exposition.

Figure 2 The flow-matching principle



In what follows, we consider the following simplified characterisation of FMNs:

1 network topology is deterministic and fixed

2 available *production* capacities $C_{p_j}(t)$ are mutually independent for all j and all t

3 network has multiple POD $(d=1,...,D)$

4 capacity propagation delays are deterministic

5 absence of backup capacity and inventory buffers.

Assumption (2) above can be justified by noting that, for large and extended supply chains, such as those of global pharmaceutical and medical equipment/devices enterprises, the various points of assembly are geographically isolated and/or operating independently. However, this assumption may not prove reasonable for disruptive events triggered by nationwide labour strikes and political unrest.

APPENDIX B (Continued)

Now, with assumptions (1)–(5) above, FMNs with multiple POD must posses the following properties.

1 $N_1 = \{1, 2, ..., D\}, \{D+1\} \in N_2, \{n\} \in N_m$.

2 $A_{ji} = 0, \forall i, j \in N$, s.t. $j < i$. Also, $\forall i, j \in N_k, A_{ji} = 0$.

3 $\forall j \in N_{k+s}, i \in N_k, A_{ji} = 0,$ where $k = 1, 2, ..., m; s \geq 2$.

4 $\forall j \in N_k, \exists$ at least one $i \in N_{k-1},$ s.t. $A_{ji} > 0,$ where $k = 1, 2, ..., m$.

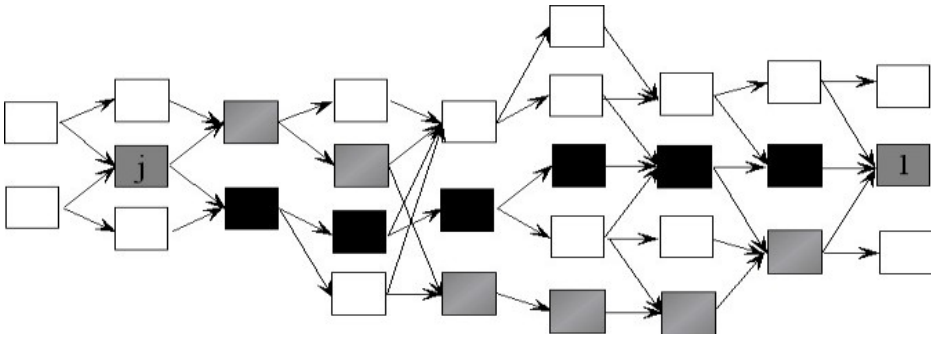
5 $\forall d \in N_1, \forall i, j \in N_2, A_{id} = A_{jd} > 0$ (equal to the propositions of the distribution of the final product among the PODs).

6 $\forall j \in N, A_{jj} = 1$ and $T_{jj} = 0$.

We now define the ith path from vertex $j \in N_k$ to vertex $l \in N_s$ as following (see Figure 3):

$$L_{jl}^i = \left\{ j, j_{k-1}^i \in N_{k-1}, j_{k-2}^i \in N_{k-2}, ..., l \in N_s : A_{j j_{k-1}^i} > 0, ..., A_{j_{s+1}^i l} > 0 \right\}.$$

Figure 3 Paths in FFN



Then, for each path L_{jl}^i, we define the product of the corresponding multifurcation coefficients as

$$A_{jl}^i = A_{j j_{k-1}^i} A_{j_{k-1}^i j_{k-2}^i} \cdots A_{j_{s+1}^i l}.$$

We let $\bar{A}_{jl} = \min_i \{A_{jl}^i\}$. In addition, for each unique path L_{jl}^i, we define T_{jl}^i as the *total capacity propagation delay of path* L_{jl}^i:

$$T_{jl}^i = T_{j j_{ik-1}} + T_{j_{k-1}^i j_{k-2}^i} + \cdots + T_{j_{s+1}^i l}.$$

We now let $\bar{T}_{jl}$ be the total capacity propagation delay of the path with the smallest value of A_{jl}^i.

In the next sections, we are concerned with characterising the probability distribution functions of individual POD as well as the overall network throughput.

APPENDIX B (Continued)

2.2 Probability law on the available effective capacity of a POD

In this section, we first derive the time-fixed distribution function of the available effective capacity of point of delivery $d(d=1,...,D)$. We then use this result in the next section to obtain the time-fixed probability law of the available effective capacity of the entire supply chain network.

We begin by noting that for a fixed $t>0$, the available supply capacity of vertex $i \in N_k$ can be expressed as

$$C_{s_i}(t) = \min_{\substack{j \in N_{k+1} \\ A_{ji}>0}} \{A_{ji} C_{ej}(t - T_{ji})\}. \tag{2}$$

Equation (2) essentially suggests that the most disrupted feeding vertex in tier N_{k+1} becomes the determinant of the available supply capacity of vertex i in tier N_k. It follows then that for vertex $i \in N_k, k < m$, and a fixed $t>0$,

$$C_{e_i}(t) = \min\{C_{p_i}(t), C_{s_i}(t)\} = \min\left\{C_{p_i}(t), \min_{\substack{j \in N_{k+1} \\ A_{ji}>0}} \left\{A_{ji} C_{e_j}(t - T_{ji})\right\}\right\}.$$

We are now in a position to express the available effective capacity of a point of delivery in the form of the following proposition.

Proposition 2.1: For a fixed time $t>0$, the available effective capacity of POD d is given by

$$C_{e_d}(t) = \min_{\substack{M=2,3,...,m \\ i_1=d, i_2 \in N_2, ..., i_m \in N_m \\ A_{i_k i_{k-1}}>0 \forall k=1,...,m}} \left\{C_{p_d}(t), C_{p_{i_M}}\left(t - \sum_{k=2}^{M} T_{i_k i_{k-1}}\right) \prod_{k=2}^{M} A_{i_k i_{k-1}}\right\}$$

$$= \min\left\{C_{p_d}(t), \min_{\substack{j \in N \setminus N_1 \\ \bar{A}_{jd}>0}} \left\{\bar{A}_{jd} C_{p_j}\left(t - \bar{T}_{jd}\right)\right\}\right\},$$

provided that $C_{e_i}(t) = C_{p_p}(t) \forall i \in N_m$ and $\forall t > 0$.

The proof of this result is presented in the Appendix.

Proposition (2.1) essentially asserts that factoring in the topology and magnitude of propagation delays, the vertex with the most disrupted available production capacity determines the available effective capacity of the corresponding point of delivery. Recall that the available production capacities of all vertices were assumed to be mutually independent at all times. We now introduce the following result.

Proposition 2.2: For a fixed time $t>0$, the complimentary distribution of the available effective capacity of POD d is given by

$$\bar{F}_{C_{e_d}(t)}(\alpha) = P\left\{C_{e_d}(t) > \alpha\right\} = \prod_{j \in E_1} \bar{F}_{C_{p_j}\left(t - \bar{T}_{jd}\right)}\left(\alpha / \bar{A}_{jd}\right), \text{ where } E_1 = \{N \setminus N_1\} \cup \{d\}.$$

The proof of this result is presented in the Appendix.

APPENDIX B (Continued)

A product form of proposition (2.2) suggests that lean FMNs can be vulnerable to even modest upstream disruptions. To a substantial extent, the degree of vulnerability is determined by the nature of topology and capacity propagation delays. In addition, relative disposition of the impacted vertex and the dynamics of the disruption itself along with the recovery dynamics will ultimately determine the overall sustainability of the POD.

2.3 Probability law on the available effective FMN capacity

Since the total available effective network capacity is the sum of the available effective capacities of the POD, the former can be expressed as following.

$$\sum_{d=1}^{D} C_{e_d}(t) = \sum_{d=1}^{D} \left[\min \left\{ C_{p_d}(t), \min_{\substack{j \in N \setminus N_1 \\ \bar{A}_{jd} > 0}} \left\{ \bar{A}_{jd} C_{p_j}\left(t - \bar{T}_{jd}\right) \right\} \right\} \right]. \tag{3}$$

At this point, we introduce two simplifying assumptions for the output tier, driven by our motivation to obtain closed form solution for the probability law on the available effective network capacity. First, we prescribe that for any point of delivery $d, C_{p_d}(t) \geq C_{s_d}(t) \forall t.$ This assumption can be found reasonable in most cases when the final assembly takes place in the USA, where the pharmaceutical and medical equipment/devices manufacturing, in general, is at least as reliable as that of the overseas suppliers, such as those located in Asia (Davidson et al., 2005). We also assume that for any point of delivery d and $\forall i, j \in N_2, T_{id} = T_{jd} = \hat{t}.$ This simplification can, in part, be justified by noting that for globally distributed supply networks with the final assembly in US, the propagation times at the last tier are small compared to the total propagation time throughout the entire network, and, thus, the former can be assumed equal.

We now consider a point of delivery d and aim to obtain an expression for the available effective capacity of d in terms of its 'most constrained supplier' $v \in N_2.$ To this end, recall that $\forall i, j \in N_2, A_{id} = A_{jd}.$ Let $h \in N_2$ and define $A_{hd} = \hat{a}_d$ and $T_{hd} = \hat{t}.$ Let

$$v = \arg \min_{h \in N_2} \left\{ C_{e_h}\left(t - \hat{t}\right) \right\}. \tag{4}$$

We are now ready to state the following important proposition.

Proposition 2.3: For a fixed time $t > 0,$ the available effective capacity of POD d is given by

$$C_{e_d}(t) = \hat{a}_d C_{e_v}\left(t - \hat{t}\right).$$

Proof: Using proposition 2.1 and provided that $C_{p_d}(t) > C_{s_d}(t)$ holds for all $t,$ we have that

APPENDIX B (Continued)

$$
\begin{aligned}
C_{e_d}(t) &= \min\left\{ C_{p_d}(t), \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jd}>0}} \left\{ \overline{A}_{jd} C_{p_j}\left(t-\overline{T}_{jd}\right)\right\}\right\} \\
&\quad \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jd}>0}} \left\{ \overline{A}_{jd} C_{p_j}\left(t-\overline{T}_{jd}\right)\right\} \\
&= \hat{a}_d \min_{h\in N_2}\left\{ \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jh}>0}} \left\{ \overline{A}_{jh} C_{p_j}\left(t-\overline{T}_{jh}-\hat{t}\right)\right\}\right\} \\
&= \hat{a}_d \min_{h\in N_2}\left\{ \min\left\{ C_{p_h}\left(t-\hat{t}\right), \min_{\substack{j\in N\setminus (N_1\cup N_2) \\ \overline{A}_{jh}>0}} \left\{ \overline{A}_{jh} C_{p_j}\left(t-\overline{T}_{jh}-\hat{t}\right)\right\}\right\}\right\} \\
&= \hat{a}_d \min_{h\in N_2}\left\{ C_{e_h}\left(t-\hat{t}\right)\right\} \\
&= \hat{a}_d C_{e_v}\left(t-\hat{t}\right).
\end{aligned}
$$

Immediately, we have the following result for expressing the available effective network capacity.

Proposition 2.4: For a fixed time $t>0$, the available effective network capacity can be expressed as

$$\sum_{d=1}^{D} C_{e_d}(t) = C_{e_v}\left(t-\hat{t}\right).$$

Proof: Using proposition 2.3,

$$\sum_{d=1}^{D} C_{e_d}(t) = \sum_{d=1}^{D} \hat{a}_d C_{e_v}\left(t-\hat{t}\right) = C_{e_v}\left(t-\hat{t}\right)\sum_{d=1}^{D}\hat{a}_d = C_{e_v}\left(t-\hat{t}\right).$$

Proposition 2.4 provides a general expression for the total available effective capacity of a FMN with multiple points a delivery in terms of the available effective capacity of some critical supplier $v\in N_2$.

We are now ready to state the main result of this section.

Proposition 2.5: For a fixed $t>0$, the complementary distribution of the available effective capacity of a FMN with multiple POD is given by the following expression, where $E_2 = \{N\setminus(N_1\cup N_2)\}\cup\{v\}$.

$$P\left\{\sum_{d=1}^{D} C_{e_d}(t) > \alpha\right\} = P\left\{C_{e_v}\left(t-\hat{t}\right) > \alpha\right\}\prod_{j\in E_2}\overline{F}_{C_{p_j}\left(t-\overline{T}_{jv}-\hat{t}\right)}\left(\alpha/\overline{A}_{jv}\right).$$

Proof: The proof combines proposition 2.4 and proposition 2.1, applied to vertex v.

APPENDIX B (Continued)

In the next section, we impose time dynamics for the available production capacity at the vertex level, subject to random disruptions, and examine the resulting capacity trajectory of the network.

3 Stochastic trajectory of available effective capacity of FMN with multiple POD

Proposition (2.5) asserts that the stochastic process $\{\sum_{d=1}^{D} C_{e_d}(t), t \geq 0\}$ is dependent, through the underlying topology, on the family of stochastic processes $\{C_{p_j}(t), t \geq 0\}$. In this part of analysis, we first discuss an instantaneous capacity loss with random recovery delay and a constant rate recovery model for the $\{C_{p_j}(t), t \geq 0\}$ family. An instantaneous capacity degradation can be used to describe a sudden failure in supporting infrastructure, such as supply of electric power, chemical agents, and cyber grid. A graceful recovery can approximate equipment warm-up, progressive maintenance, or stepwise modernisation. Another example of such sample path is a terrorist attack warning causing an immediate assembly shutdown, followed by area check-up and gradual release of production capacity.

3.1 Instantaneous capacity loss with random recovery delay and constant recovery rate model

Consider the following dynamics of the available production capacity of vertex j. The system operates at the target capacity level C^*. Disruptions occur one at a time, each rendering an instantaneous loss, followed by a random delay and a constant rate $\alpha > 0$ recovery (Figure 4).

The points of recovery form a sequence of stopping times, at which the process stochastically regenerates. Let $X_n, n \in \mathbb{N}$ be the amount of time the vertex operates at the C^* level before the n-th shock. $\{X_n, n \in \mathbb{N}\}$ are assumed to be i.i.d. random variables with mean μ_X. Let ΔC_n be the magnitude of the n-th loss $(0 \leq \Delta C_n \leq C^*, n \in \mathbb{N})$, and assume $\{\Delta C_n, n \in \mathbb{N}\}$ are i.i.d. random variables with mean $\mu_{\Delta C}$, independent of the $\{X_n\}$. Let $R_n, n \in \mathbb{N}$ be a random recovery delay during which the available production capacity remains at level $C^* - \Delta C_n$ (Figure 4). We assume that $\{R_n\}$ form a sequence of i.i.d. random variables with mean μ_R, and mutually independent of $\{X_n\}$ and ΔC_n. We let $Y_n, n \in \mathbb{N}$ be the length of the n-th recovery period

$$Y_n = \frac{\Delta C_n}{\alpha}, \alpha > 0.$$

Let $Z_n = \sum_{i=1}^{n}(X_i + R_i + Y_i), n \in \mathbb{N}$, whereby $\{Z_n\}$ forms an embedded renewal process. We first seek to derive the limiting complimentary distribution $\lim_{t\to\infty} P\{C_{p_j}(t) \geq u\}$. Beginning with the $\{C_{p_j}(t), t \geq 0\}$ process, we define a regenerative process $\{B^u(t), 0 \leq u \leq C^*, t \geq 0\}$, with state space {0, 1} where

$$B^u(t) = \begin{cases} 1, & \text{if } C_{p_j}(t) \geq u; \\ 0, & \text{otherwise.} \end{cases}$$

APPENDIX B (Continued)

The process is in state '1' when the amount of the available production capacity is at least u, and it is in state '0' otherwise. At the epoch, where a disruption causes the available production capacity to fall below u, a transition occurs from state '1' to state '0'. At the next epoch, where the available production capacity recovers to level u, a transition occurs from state '0' to state '1', and the process regenerates. A cycle is defined as the time period between occurrence of two successive transitions from '0' to '1' (see Figure 4). Clearly, the $\{B^u(t), t \geq 0\}$ process has the same stopping times as the process $\{C_{p_j}(t), t \geq 0\}$.

Consider the first cycle of the process $\{B^u(t), t \geq 0\}$ and let T_u denote the portion of the cycle during which the process is in state '1' [i.e., the value of T_u exceeds u, (Figure 4)]. Let T denote the length of the cycle, and let the index of the (first) capacity loss that causes the process to make a transition from state '1' to state '0' be as follows

$$N_u = n \text{ s.t. } \Delta C_n > C^* - u. \tag{5}$$

It then follows that

$$T = \frac{C^* - u}{\alpha} + Z_{N_u} - \frac{C^* - u}{\alpha} = Z_{N_u}, \tag{6}$$

and

$$T_u = \frac{C^* - u}{\alpha} + Z_{N_u} - \left(R_{N_u} + \frac{\Delta C_{N_u}}{\alpha} \right). \tag{7}$$

We then let

$$p_u = P\{\Delta C_n > C^* - u\} \text{ for } n = 1, 2, ..., N_u.$$

Note that N_u has a negative binomial distribution with the probability mass function given by

$$P\{N_u = n\} - p_u(1 - p_u)^{(n-1)} \text{ for } n = 1, 2, ... \tag{8}$$

We are now in a position to state the following important result.

Proposition 3.1: For the instantaneous capacity loss with random recovery delay and a constant recovery rate model, the limiting complimentary distribution of the available production capacity of vertex j is given by the following expression, where $E(N_u) = \sum_{k=1}^{\infty} k p_u (1 - p_u)^{(k-1)}, 0 \leq u \leq C^*$:

$$\lim_{t \to \infty} P\{C_{p_j}(t) \geq u\} = \frac{(C^* - u) + E(N_u)(\alpha\mu_R + \mu_{\Delta C}) - \alpha\mu_R - \mu_{\Delta C}}{E(N_u)(\alpha\mu_X + \alpha\mu_R + \mu_{\Delta C})}.$$

Proof: From equations (6)–(8), for this alternating regenerative stochastic process we observe that

$$\lim_{t \to \infty} P\left\{C_{p_j}(t) \geq u\right\} = \frac{E(T_u)}{E(T)}$$
$$\frac{(C^* - u) + E(N_u)(\alpha\mu_X + \alpha\mu_R + \mu_{\Delta C}) - \alpha\mu_R - \mu_{\Delta C}}{E(N_u)(\alpha\mu_X + \alpha\mu_R + \mu_{\Delta C})}.$$

APPENDIX B (Continued)

We have the following important corollary when $u = C^*$.

Corollary 3.1: For the instantaneous capacity loss with random recovery delay and a constant recovery rate model, the limiting probability that vertex j operates at the target capacity level is given by

$$\lim_{t\to\infty} P\left\{C_{p_j}(t) = C^*\right\} = \frac{\mu_X}{\mu_X - \mu_R + (\mu_{\Delta C}/\alpha)}.$$

Proof: Note that in this case, $p_u = 1$ and $E(N_u) = 1$. The result follows from proposition (3.1). □

Despite its seemingly simple form, corollary 3.1 can be used for sensitivity analysis-based decision support, as shown, e.g., in Figure 5 and Figure 6.

Figure 5 Sensitivity on $\mu_{\Delta C}$ and μ_X

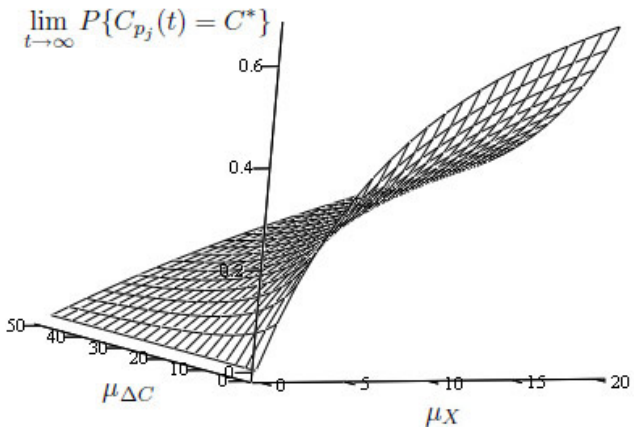
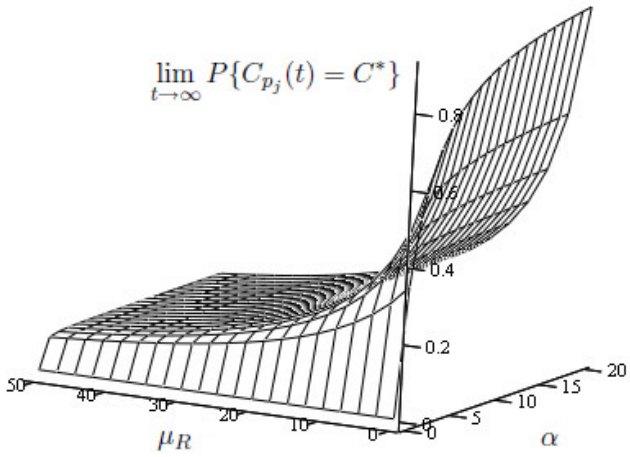


Figure 6 Sensitivity on μ_R and α



APPENDIX B (Continued)

As shown in Figure 5, there could be a certain trade-off in attempting to control the rate and magnitude of disruptions (via μ_X and $\mu_{\Delta C}$, respectively, assuming the other variables are fixed). Increasing the value of μ_X results in the limiting target capacity probability attaining its asymptomatic level in a steep or more gradual fashion, depending on the mean magnitude of disruption $\mu_{\Delta C}$. On the other hand, if a decision-maker has no control on the disruption dynamics, she may instead prefer a mitigation strategy focused on agile recovery by effecting the mean recovery delay μ_R and recovery rate α. As can be seen in Figure 6, when the other variables are kept fixed, for small values of μ_R, the recovery rate can have a significant effect on the resultant limiting probability. On the other hand, as recovery delays become more prolonged (particularly for comparable values of μ_X), attempting a more rapid recovery can yield only marginal impact on the limiting probability, which reaches saturation at lower levels.

We conclude this section with the following theorem providing the desired expression for the limiting distribution of the available effective FMN capacity.

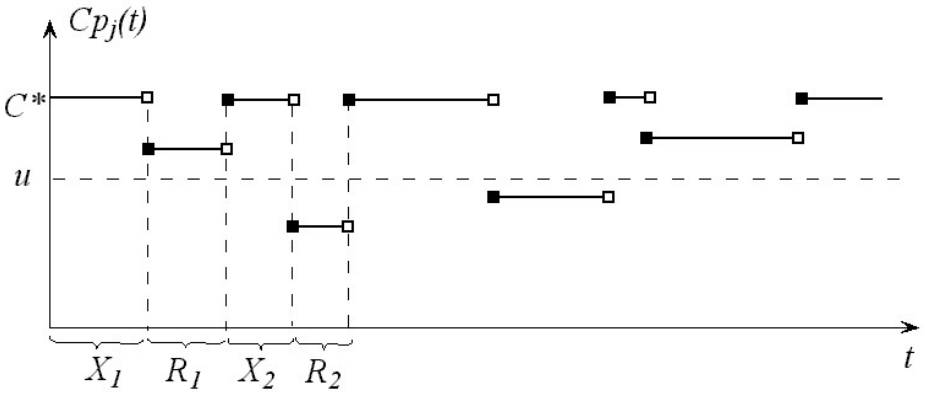
Theorem 3.2: For the instantaneous capacity loss with random recovery delay and a constant recovery rate model, the limiting complimentary distribution of the available effective FMN capacity is given by the following, where $\tilde{u}_{jv} = u / \overline{A}_{jv}$ and $E_2 = \{N / (N_1 \cup N_2)\} \cup \{v\}$ [see also equation (4)]:

$$\lim_{t\to\infty} P\left\{\sum_{d=1}^{D} C_{e_d}(t) \geq u\right\} = \prod_{j\in E_2} \frac{(C^* - \tilde{u}_{jv}) / \alpha + | E\left(N_{\tilde{u}_{jv}}\right)(\alpha\mu_X + \alpha\mu_R + \mu_{\Delta C}) - \alpha\mu_R - \mu_{\Delta C}}{E\left(N_{\tilde{u}_{jv}}\right)(\alpha\mu_X + \alpha\mu_R + \mu_{\Delta C})}.$$

Proof: Combine propositions (2.5) and (3.1).□

As shown next in Section 3.2 and further discussed in Section 4, the model discussed above can serve as a foundation for a family of predictive models attempting to describe propagation of stochastic dynamics of capacity disruptions at the vertex level, to examine the limiting behaviour of the network.

Figure 7 Realisation of instantaneous capacity loss with random recovery delay and instantaneous restoration



APPENDIX B (Continued)

3.2 Instantaneous capacity loss with random recovery delay and instantaneous recovery

In this section, we present a rather straightforward special case of the model discussed in Section 3.1. When recovery rate $\alpha = \infty$, the model exhibits an instantaneous capacity loss followed by a random recovery delay and instantaneous restoration, as shown in Figure 7. The simplicity of this basic model can be appreciated by practitioners in the cases when capacity ramp-up times Y_i are small compared to X_i and R_i.

In this section, we state the results without proofs which are trivial by substituting $\alpha = \infty$ into the results of the previous section.

Proposition 3.3: For the instantaneous capacity loss with random recovery delay and instantaneous recovery model, the limiting complimentary distribution of the available production capacity of vertex j is given by the following expression, where $E(N_u) = \sum_{k=1}^{\infty} k p_u (1-p_u)^{(k-1)}, 0 \le u \le C^*$:

$$\lim_{t\to\infty} P\{C_{p_j}(t) \ge u\} = 1 - \frac{\mu_R}{E(N_u)(\mu_X + \mu_R)}.$$

Corollary 3.2: For the instantaneous capacity loss with random recovery delay and instantaneous recovery model, the limiting probability that vertex j operates at the target capacity is given by

$$\lim_{t\to\infty} P\{C_{p_j}(t) = C^*\} = \frac{\mu_X}{\mu_X + \mu_R}.$$

Finally, by combining propositions (2.5) and (3.3), we have the limiting distribution of the available effective FMN capacity in the form of the following theorem.

Theorem 3.4: For the instantaneous capacity loss with random recovery delay and instantaneous recovery model, the limiting complimentary distribution of the available effective FMN capacity is given by the following, where $\tilde{u}_{jv} = u / \overline{A}_{jv}$ and $E_2 = \{N \setminus (N_1 \cup N_2)\} \cup \{v\}$ [see also equation (4)]:

$$\lim_{t\to\infty} P\left\{\sum_{d=1}^{D} C_{e_{jd}}(t) \ge u\right\} = \prod_{j\in E_2} \frac{E(N_{\tilde{u}_{jv}})(\mu_X + \mu_R) - \mu_R}{E(N_{\tilde{u}_{jv}})(\mu_X + \mu_R)}.$$

In the next section, we discuss how the formalism developed in Section 3 can provide a foundation for developing future scalable, risk-driven methodologies for analytical and computational support of strategic design of healthcare supply chain systems.

4 Conclusions

Reengineering the USA existing healthcare delivery into a timely and efficient system has to invariably rely on robust and agile supply networks. Current healthcare supply chain systems, such as pharmaceutical and medical equipment/devices enterprises, can span multiple continents and subsume hundreds of suppliers and customers. In addition, as these enterprises have been adopting the philosophy of lean manufacturing, including

APPENDIX B (Continued)

slimming inventory buffers, global outsourcing, and consolidating supplier base, such reductionism has also left them operating in an increasingly risk-encumbered environment. One of the most profound risk factor is unpredicted disruptions in available capacity, as might be caused by a number of controllable and uncontrollable factors, including nature, process hazards, and human activity. At the same time, efficient predictive analytics and computational tools suitable for analysing the impact of capacity disruptions on network risk have been largely unavailable.

In this paper, we present an original analysis of healthcare supply chain systems exhibiting converging assembly and exposed to random capacity disruptions. The supply chain was modelled as a feed-forward FMN with multiple POD – a mathematical formalism particularly useful for understanding changes in the aggregate network capacity, as might be impacted by unpredicted perturbations. In our novel work, we captured the time-fixed probability law on the available effective network capacity, in the presence of capacity propagation delays, in closed form. We then construed two models of stochastic dynamics of the available effective capacity at the network level. As discussed in Section 3, these trajectories can be used to model a number of disruptive scenarios.

Our analysis will contribute to understanding the degree of risk exposure and vulnerability of global healthcare supply chains. Moreover, our study can be used to provide a substantive measure of the trade-off between a lean structure of the supply chain and its robustness and agility. The stochastic models of capacity disruptions discussed in this paper present one of the initial attempts to characterise supply chain dynamics via a FMN-based formalism. In the future, these stochastic models of capacity dynamics will be generalised to feature different recovery modes and incorporate disposition of strategic inventory buffers or capacity back-ups. This stochastic analytics will then be combined with the dynamics of demand at the POD and capacity expenditures, to develop future methodologies, based on the tenets of utility theory, to provide decision support for design of resilient healthcare supply chain systems.

Acknowledgements

The authors would like to thank the referees and the editor for their help to improve the quality of the paper. This work was supported by National Science Foundation Grant CMMI 0621030.

References

Abboud, N.E. (1997) ‘A simple approximation of the EMQ model with Poisson machine failures’, *Production Planning and Control*, Vol. 8, pp.385–397.

Arcelus, F.J. and Srinivasan, G. (1995) ‘Discount strategies for one-time only sales’, *IIE Transactions*, Vol. 27, pp.618–624.

Ardalan, A. (1995) ‘A comparative analysis of approaches for determining optimal price and order quantity when a sale increases demand’, *European Journal of Operational Research*, Vol. 84, pp.416–430.

Arreola-Risa, A. and DeCroix, C.A. (1998) ‘Inventory management under random supply disruptions and partial backorders’, *Naval Research Logistics*, Vol. 45, pp.687–703.

APPENDIX B (Continued)

Atamt, A. and Zhang, M. (2007) 'Two-stage robust network flow and design under demand uncertainty', *Operations Research*, Vol. 55, pp.662–673.

Aull-Hyde, R.L. (1992) 'Evaluation of supplier-restricted purchasing options under temporary price discounts', *IIE Transactions*, Vol. 24, pp.184–186.

Bayer, S., Koberle-Gaiser, M. and Barlow, J. (2007) 'Planning for adaptability in healthcare infrastructure', *Proceedings of the 25th International System Dynamics*, 29 July–2 August, p.978, Boston, USA.

Bielecki, R.E. and Kumar, P.R. (1988) 'Optimality of zero inventory policies for unreliable manufacturing systems', *Operations Research*, Vol. 36, pp.532–541.

Booth, R. (1996) 'The role of supply-chain re-engineering in the pharmaceutical industry', *Logistics Information Management*, Vol. 9, pp.4–10.

Brennan, C. (1998) 'Integrating the healthcare supply chain', *Healthcare Financial Management: Journal of the Healthcare Financial Management Association*, Vol. 52, No. 1, p.31.

Buzacott, J.A. and Shantikumar, J.G. (1993) *Stochastic Models of Manufacturing Systems*, Prentice Hall, Englewood Cliffs, NJ.

Chao, H. (1987) 'Inventory policy in the presence of market disruptions', *Operations Research*, Vol. 2, pp.274–281.

Coons, S. (2005) 'After Katrina: will lessons be learned and progress made?', *Clinical Therapeutics*, Vol. 27, No. 10, pp.1620–1621.

Davidson, L., Greblov, G. and Bloomington, I. (2005) 'The pharmaceutical industry in the global economy', Indiana University Kelley School of Business Bloomington, Indiana, available at http://www. bus. indiana.edu/davidso/lifesciences/lsresearchpapers/pharmaceutical% 20indust ryaug12.doc (accessed on 16 August 2006).

Dorogovtsev, S. and Mendes, J. (2003) *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford, Great Britain.

Eiselt, H., Gendreau, M. and Laporte, G. (1996) 'Optimal location of facilities on a network with an unreliable node or link', *Information Processing Letters*, Vol. 58, pp.71–74.

Ertas, A., Tanik, M. and Maxwell, T. (2000) 'Transdisciplinary engineering education and research model', *Journal of Integrated Design and Process Science*, Vol. 4, No. 4, pp.1–11.

Gallego, G. (1988a) 'Linear control policies for scheduling a single facility after an initial disruption', Technical report 770, School of OR and IE, Cornell University, Ithaca, NY.

Gallego, G. (1988b) 'Produce-up-to policies for scheduling a single facility after an initial disruption' Technical report 771, School of OR and IE, Cornell University, Ithaca, NY.

Groenevelt, H., Seidmann, A. and Pintelon, L. (1992) 'Production lot sizing with machine breakdowns', *Management Science*, Vol. 38, pp.104–123.

Hall, P. (2004) 'We'd have to sink the ships: impact studies and the 2002 West Coast Port lockout', *Economic Development Quarterly*, Vol. 18, No. 4, pp.354–367.

Harland, C. (1996) 'Supply network strategies the case of health supplies', *European Journal of Purchasing and Supply Management*, Vol. 2, No. 4, pp.183–192.

Henig, M. and Gerchak, Y. (1990) 'The structure of periodic review policies in the presence of random yield', *Operations Research*, Vol. 38, pp.634–643.

Hopp, W.J., Pati, N. and Jones, P.C. (1989) 'Optimal inventory control in a production flow system with failures', *International Journal of Production Research*, Vol. 27, pp.1367–1384.

Horwich, G. (2000) 'Economic lessons of the Kobe earthquake', *Economic Development and Cultural Change*, Vol. 48, No. 3, pp.521–542.

Hurtado, M., Corrigan, J. and Swift, E. (2001) *Envisioning the National Health Care Quality Report*, National Academy Press.

Institute of Medicine (2001) *Crossing the Quality Chasm: A New Health System for the 21st Century*, National Academy Press Washington, DC.

APPENDIX B (Continued)

Isidore, C. (2002) 'Hope in West Coast Port talks', available at http://money.cnn.com/2002/10/02/news/economy/ports/index.htm.

Kim, C.H. and Hong, Y. (1997) 'An extended EMQ model for a failure prone machine with general lifetime distribution', *International Journal of Production Economics*, Vol. 49, pp.215–223.

Koh, R., Schuster, E., Chackrabarti, I. and Bellman, A. (2003) 'Securing the pharmaceutical supply chain', Auto-ID Center MIT, White Paper.

Lee, H.L. (1992) 'Lot sizing to reduce capacity utilization in a production process with defective items, process corrections, and rework', *Management Science*, Vol. 38, pp.1314–1328.

Lee, S. (2001) 'On solving unreliable planar location problems', *Computers and Operations Research*, Vol. 28, pp.329–344.

Lemos, R. (2003a) 'Slammer' attacks may become way of life for net', available at http://news.com.com/Damage+control/2009-1001_3-983540.html.

Lemos, R. (2003b) 'Worm exposes apathy, Microsoft flaws', available at http://news.com.com/2100-1001-982135.html.

M2 Comm. (2005) 'Asiana airlines apologizes for strike disruption', available at http://www.allbusiness.com.

Meyer, R.R., Rothkpof, M.H. and Smith, S.A. (1979) 'Reliability and inventory in a production storage system', *Management Science*, Vol. 25, pp.799–807.

Mohd, N. (2006) *Supply Chain Risk Management: A Framework*, SCMS-COCHIN.

Mohebbi, E. (2003) 'Supply interruptions in a lost-sales inventory system with random lead time', *Computers and Operations Research*, Vol. 30, pp.824–835.

Papageorgiou, L., Rotstein, G. and Shah, N. (2001) 'Strategic supply chain optimization for the pharmaceutical industries', *Ind. Eng. Chem. Res*., Vol. 40, No. 1, pp.275–286.

Parlar, M. (1997) 'Continuous-review inventory problem with random supply interruptions', *European Journal of Operational Research*, Vol. 99, pp.366–385.

Parlar, M. and Berkin, D. (1991) 'Future supply uncertainty in EOQ models', *Naval Research Logistics*, Vol. 38, pp.50–55.

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

Petersen, K., Handfield, R. and Ragatz, G. (2005) 'Supplier integration into new product development: coordinating product, process and supply chain design', *Journal of Operations Management*, Vol. 23, Nos. 3–4, pp.371–388.

Porter, M. and Teisberg, E. (2004) 'Redefining competition in health care', *Harvard Business Review*, pp.64–77.

Porteus, E.L. (1986) 'Optimal lot sizing, process quality improvement and setup cost reduction', *Operations Research*, Vol. 34, pp.137–144.

Posner, M.J. and Berg, M. (1989) 'Analysis of a production-inventory system with unreliable production facility', *Operations Research Letters*, Vol. 8, pp.339–345.

RNCOS (2008) 'Global pharmaceutical market forecast to 2012', available at http://www.prlog.org/10124036-global-pharmaceutical-market-forecast-to-2012.html.

Rosenbaum, S. (2006) 'US health policy in the aftermath of Hurricane Katrina', *Jama*, Vol. 295, No. 4, pp.437–440.

Rosenblatt, M.J. and Lee, H.L. (1986) 'Economic production cycles with imperfect production processes', *IIE Transactions*, Vol. 18, pp.48–55.

Sarkisian, M. (1995) 'Cost lessons of the Great Hanshin earthquake', *Cost Engineering*, Vol. 37, No. 8, pp.13–16.

Savachkin, A., Bakir, N.O. and Uribe-Sanchez, A. (2008) 'An optimal countermeasure policy to mitigate random capacity disruptions in a production system', *International Journal of Agile Systems and Management*, Vol. 3, Nos. 1/2, pp.4–17.

APPENDIX B (Continued)

Shah, N. (2004) 'Pharmaceutical supply chains: key issues and strategies for optimization', *Computers and Chemical Engineering*, Vol. 28, Nos. 6–7, pp.929–941.

She, Y. (2001) 'Supply chain management under the threat of international terrorism', *International Journal of Logistics Management*, Vol. 12, No. 2, pp.1–12.

Snyder and Daskin (2005) 'Reliability models for facility location: the expected failure cost case', *Transportation Science*, Vol. 39, pp.400–416.

Snyder, Daskin and Teo (2007) 'The stochastic location model with risk pooling', *European Journal of Operational Research*, Vol. 179, pp.1221–1238.

Song, J. and Zipkin, P. (1996) 'Inventory control with information about supply conditions', *Management Science*, Vol. 42, pp.1409–1419.

Taylor, S.G. and Bradley, C.E. (1985) 'Optimal ordering strategies for announced price increases', *Operations Research*, Vol. 33, pp.312–325.

Tersine, R.J. and Barman, S. (1995) 'Economic purchasing strategies for temporary price discounts', *European Journal of Operational Research*, Vol. 80, pp.328–343.

Tsiakis, P., Shah, N. and Pantelides, C. (2001) 'Design of multi-echelon supply chain networks under demand uncertainty', *Industrial and Engineering Chemistry Research*, Vol. 40, pp.3585–3604.

Weiss, H.J. and Rosenthal, E.C. (1992) 'Optimal ordering policies when anticipating a disruption in supply or demand', *European Journal of Operational Research*, Vol. 59, pp.370–382.

Appendix

A1 Proof of proposition 2.1

Condition $C_{e_i}(t) = C_{p_i}(t) \forall\, i \in N_m$ means that all vertices in the input tier N_m have no suppliers, i.e., they represent a starting point of assembly. Recursively, we have that

$$
\begin{aligned}
C_{e_d}(t) &= \min\left\{C_{p_d}(t), C_{s_d}(t)\right\} \\
&= \min_{i_2 \in N_2}\left\{C_{p_d}(t), A_{i_2 d} C_{e_{i_2}}\left(t - T_{i_2 d}\right)\right\} \\
&= \min_{i_2 \in N_2}\left\{C_{p_d}(t), \min_{\substack{i_3 \in N_3 \\ A_{i_3 i_2} > 0}}\left\{A_{i_2 d} C_{p_{i_2}}\left(t - T_{i_2 d}\right), A_{i_2 d} A_{i_3 i_2} C_{e_{i_3}}\left(t - T_{i_2 d} - T_{i_3 i_2}\right)\right\}\right\} \\
&= \min_{\substack{i_2 \in N_2, i_3 \in N_3 \\ A_{i_3 i_2} > 0}}\left\{C_{p_d}(t), A_{i_2 d} C_{p_{i_2}}\left(t - T_{i_2 d}\right), A_{i_3 i_2} A_{i_2 d} C_{e_{i_3}}\left(t - T_{i_2 d} - T_{i_3 i_2}\right)\right\} \\
&\cdots \\
&= \min_{\substack{M = 2,3,\ldots,m \\ i_1 = d, i_2 \in N_2, \ldots, i_m \in N_m \\ A_{i_k i_{k-1}} > 0 \forall\, k = 1,\ldots,m}}\left\{C_{p_d}(t), C_{p_{i_M}}\left(t - \sum_{k=2}^{M} T_{i_k i_{k-1}}\right)\prod_{k=2}^{M} A_{i_k i_{k-1}}\right\}.
\end{aligned}
$$

Then we have the following.

APPENDIX B (Continued)

$$C_{e_d}(t) = \min_{\substack{M=2,3,\ldots,m \\ i_1=d, i_2\in N_2,\ldots,i_m\in N_m \\ A_{i_k i_{k-1}}>0 \forall k=1,\ldots,m}} \left\{ C_{p_d}(t), C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right)\prod_{k=2}^{M} A_{i_k i_{k-1}} \right\}$$

$$= \min_{M=2,3,\ldots,m} \left\{ C_{p_d}(t), C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right) \min_{\substack{d=1, i_2\in N_2,\ldots,i_m\in N_m \\ A_{i_k i_{k-1}}>0 \forall k=1,\ldots,m}} \left\{ \prod_{k=2}^{M} A_{i_k i_{k-1}} \right\} \right\}.$$

For each fixed vertex i_M, the expression

$$C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right)\prod_{k=2}^{M} A_{i_k i_{k-1}},$$

where $i_1 \in N_1, i_2 \in N_2, \ldots, i_m \in N_m; A_{i_k i_{k-1}} > 0,$ is equivalent to

$$C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right) A^{i}_{i_M d}$$

for some path i from vertex i_M to POD d. Then

$$C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right) \min_{\substack{d=1, i_2\in N_2,\ldots,i_m\in N_m \\ A_{i_k i_{k-1}}>0 \forall k=1,\ldots,m}} \left\{ \prod_{k=2}^{M} A_{i_k i_{k-1}} \right\}$$

$$= C_{p_{i_M}}\left(t-\sum_{k=2}^{M} T_{i_k i_{k-1}}\right) \min_{i} \left\{ A^{i}_{i_M d} \right\} = C_{p_{i_M}}\left(t-\overline{T}_{i_M d}\right)\overline{A}_{i_M d}.$$

Finally, the available effective capacity of the point of delivery can be expressed as

$$C_{e_d}(t) = \min\left\{ C_{p_d}(t), \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jd}>0}} \left\{ \overline{A}_{jd} C_{p_j}\left(t-\overline{T}_{jd}\right) \right\} \right\}.$$

A2 Proof of proposition 2.2

From proposition 2.1 and using the independence assumption, we have that

$$\overline{F}_{C_{e_d}(t)}(\alpha) = P\left(C_{e_d}(t) > \alpha\right)$$

$$= P\left\{ \min\left\{ C_{p_d}(t), \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jd}>0}} \left\{ \overline{A}_{jd} C_{p_j}\left(t-\overline{T}_{jd}\right) \right\} \right\} > \alpha \right\}$$

$$= P\left\{ C_{p_d}(t) > \alpha \min_{\substack{j\in N\setminus N_1 \\ \overline{A}_{jd}>0}} \left\{ \overline{A}_{jd} C_{p_j}\left(t-\overline{T}_{jd}\right) \right\} > \alpha \right\}$$

$$= \prod_{j\in E_1} \overline{F}_{C_{p_j}\left(t-\overline{T}_{jd}\right)}\left(\alpha / \overline{A}_{jd}\right), \text{ where } E_1 = \{N\setminus N_1\}\cup\{d\}.$$

APPENDIX C:

PUBLICATION 2: TWO COUNTERMEASURE STRATEGIES TO MITIGATE RANDOM DISRUPTIONS IN CAPACITATED SYSTEMS

In this appendix, we present the final version of the manuscript "Two Countermeasure Strategies to Mitigate Random Disruptions in Capacitated Systems" published in the Journal of Systems Science and Systems Engineering, volume 19, number 2, pages 210-226, 2010, by Springer Publisher. The co-authors, Dr. Niyazi Bakir and Dr. Alex Savachkin, authorized to include this document in my dissertation. Springer Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in my Ph.D. dissertation is attached in Appendix A.

APPENDIX C: (Continued)

J Syst Sci Syst Eng (Jun 2010) 19(2): 210-226
DOI: 10.1007/s11518-010-5125-y
ISSN: 1004-3756 (Paper) 1861-9576 (Online)
CN11-2983/N

TWO COUNTERMEASURE STRATEGIES TO MITIGATE RANDOM DISRUPTIONS IN CAPACITATED SYSTEMS*

Niyazi O. BAKIR [1] Alex SAVACHKIN[2] Andrés Uribe-SANCHEZ[3]

[1]*Department of Industrial Engineering, Bilkent University, Turkey*

nonur@bilkent.edu.tr

[2,3] *Department of Industrial and Management Systems Engineering, University of South Florida, USA*

[2] *savachki@eng.usf.edu* (✉)

[3] *auribe@mail.usf.edu*

Abstract

We examine a capacitated system exposed to random stepwise capacity disruptions with exponentially distributed interarrival times and uniformly distributed magnitudes. We explore two countermeasure policies for a risk-neutral decision maker who seeks to maximize the long-run average reward. A one-phase policy considers implementation of countermeasures throughout the entirety of a disruption cycle. The results of this analysis form a basis for a two-phase model which implements countermeasures during only a fraction of a disruption cycle. We present an extensive numerical analysis as well as a sensitivity study on the fluctuations of some system parameter values.

Keywords: Capacity analysis, capacitated system, lean, random disruptions, countermeasure policy

1. Introduction and Motivation

Lean manufacturing philosophy and associated business practices have been widely embraced and deployed by global enterprises. Some estimates assert that the shift to JIT scheduling in the US automotive industry has saved companies more than $1 billion a year in inventory costs, alone. While lean manufacturing has substantially boosted operational efficiency, it has also left enterprises operating in an increasingly risk-encumbered environment. Capacity disruptions triggered by forces of nature, property- and process-related hazards, and man-made interventions have proven to be the most profound influence on enterprise risk. As evidenced in 1995, an earthquake hit the port town of Kobe, Japan, razed to the ground 100,000 buildings and shut down Japan's largest port for over two years. In 1999, an earthquake in Taiwan displaced power lines to the semiconductor fabrication facilities responsible for more than 50 percent of the worldwide supplies of certain computer components, and shaved 5 percent off earnings for major hardware manufacturers including Dell, Apple, Hewlett-Packard, IBM, and

* This work was supported by U.S. National Science Foundation Grant CMMI 0621030.

APPENDIX C: (Continued)

Compaq (Wilcox 1999). In September 2002, longshoremen on the US West Coast were locked out in a labor strike for 11 days, forcing the shutdown of 29 ports. With more than $300 billion of dollars in goods shipped annually through these ports, the dispute caused between $11 and $22 billion in lost sales, spoiled perishables and underutilized capacity (Isidore 2002). In December 2002, a political strike in Venezuela made transnational businesses including GM, BP, Ford, Goodyear and Procter & Gamble halt their manufacturing for the duration of the conflict (Wilson 2003). The recent 2003 outbreak of SARS in China and Singapore forced Motorola to close several plants (Berniker 2003). Man-made disasters are on the rise, from terrorist attacks to computer viruses (Lemos 2003). As a result of the above events, according to a recent survey by A.M. Best Company, Inc. of 600 executives, 69 percent of chief financial officers, treasurers and risk managers at Global 1,000 companies in North America and Europe view property-related hazards-such as fires and explosions--and supply chain disruptions as the leading threats to top revenue sources (A.M. Best Company 2006).

Historically, enterprises have lacked appropriate decision support methodologies and computational tools suitable for addressing risk incurred through capacity disruptions. In academia, traditional research efforts on minimizing the cost of supply chain operations and the focus on leveraging economies of scale often yield results that overconcentrate resources. Such optimal solutions can be very sensitive to parameter fluctuations, caused by supply chain disruptions. The inability to recognize the hidden costs of such overconcentration heightens the risk of increased costs and capacity imbalance. Much of the recent literature focuses on minimizing costs of supply chain operations (see, for example, Barness-Shuster et al. (2002), Cheung & Lee (2002), Milner & Kouvelis (2002), Corbett & DeCroix (2001), Lee et al. (1997)), whereas only a small fraction of the efforts have been dedicated to modeling the impact of various disruptions, such as those affecting demand patterns, supplier and production lead times, prices, imperfect process quality, process yield, and other factors.

One of the most common types of disruption appearing in the literature is that of supply rate changes. An excellent work by Arreola-Risa & DeCroix (1998) explores inventory management of stochastic demand systems, where the product supply is disrupted for periods of random duration. The classic economic order quantity (EOQ) problem with supply disruptions is studied by Parlar & Berkin (1991) and Parlar & Perry (1996) consider a order-quantity/ reorder-point inventory models with two suppliers subject to independent disruptions to compute the exact form of the average cost expression. Mohebbi (2003) presents an analytical model for computing the stationary distribution of the on-hand inventory in a continuous-review inventory system with compound Poisson demand, Erlang distributed lead time, and lost sales, where the supplier can assume one of the two “available” and “unavailable” states at any point in time according to a continuous-time Markov chain. Papers addressing both supply disruptions and random demand include (Chao 1987, Parlar 1997, Song & Zipkin 1996). Chao (1987)

APPENDIX C: (Continued)

proposes a dynamic model concerning optimal inventory policies in the presence of market disruptions, which are often characterized by events with uncertain arrival time, severity and duration. Parlar (1997) considers a continuous-review stochastic inventory problem with random demand and random lead-time where supply may be disrupted due to machine breakdowns, strikes or other randomly occurring events. Song & Zipkin (1996), explore an inventory-control model which includes a detailed Markovian model of the resupply system. A number of papers which address supply and demand changes have been developed in the field of oil stockpiling, as there has been grave concern over the oil supply from the Middle East (Teisberg 1981, Chap & Manne 1982, Murphy et al. 1987). Modeling production rate disruptions (machine failures) has been largely addressed by extending classical economic manufacturing quantity (EMQ) models. Rosenblatt & Lee (1986) derive an EMQ model when the production process is subject to a random deterioration from an in-control state to an out-of control state. Lee (1992) models the defect-generating process in the semiconductor wafer probe process to determine an optimal lot size, which reduces the average processing time on a critical resource. Abboud (1997) presents a simple approximation of the EMQ model with Poisson machine breakdowns and low failure rate. Groenevelt et al. (1992) study an unreliable production system with constant demand and random breakdowns, with the focus on the effects of machine failure and repair on optimal lot-sizing decisions. Assuming exponentially distributed time between failures and instantaneous repair of the machine, authors derive some unique properties of their model compared to the classical EMQ model. Groenevelt et al. (1992) extend their earlier work in Groenevelt et al. (1992) to the case where repair times are randomly distributed and excess demand is lost. Kim & Hong (1997) propose an extension to the model in Groenevelt et al. (1992), which determines an optimal lot size when a machine is subject to random failures and the time to repair is constant. They formulate average cost functions for the optimal lot size, and derive conditions for determining the optimal lot size. Hopp et al. (1989) presents a model that assumes the (*s, S*) control policy. With Poisson failures and exponential repair times, a cost function is derived. Rahim (1994) presents an integrated model for determining an economic manufacturing quantity, inspection schedule and control chart design of an imperfect production process, where he assumes that the process is subject to the occurrence of a non-Markovian shock having an increasing failure rate. Among other notable examples of such works are Henig & Gerchak (1990), Bielecki & Kumar (1988), Buzacott & Shantikumar (1993). Finally, Abboud (2001) examines a single machine production and inventory system with a deterministic production and demand rate, when the machine is subject to random failures. The author models the production/inventory system as a Markov chain and develops an algorithm to compute the potentials that are used to formulate the cost function.

At this point, we can summarize that research efforts addressing the disruption of supply are still comparatively new and scant. Most of the open literature considering various

APPENDIX C: (Continued)

types of disruptions focuses on issues of inventory, ordering, production lot sizing, production scheduling, and cost management of inventory, setup, and backorder costs. To the best of our knowledge, there have been no attempts to consider introducing countermeasure policies for mitigating unpredicted capacity disruptions in a capacitated system, and analyze the benefits of such policies for the system manager. Our paper presents an initial attempt to fill the vacuum in this area.

The paper has the following organization. In Section 2, we introduce notation and problem definition. Section 3 presents analysis of a one-phase countermeasure policy, where a risk-neutral decision maker implements countermeasures during the entirety of a disruption cycle, striving to maximize the long-run average reward. These results are used in Section 4 to examine a richer class of policies, where countermeasures are activated during only a fraction of a disruption cycle. In Section 5, we present a numerical analysis for determining the optimal phase threshold and examine the sensitivity of the optimal policy to fluctuations in system parameter values. Finally, Section 6 offers concluding remarks.

2. Notation and Problem Definition

For the rest of this paper, we define *throughput* as the long-run average of the number of item units per unit time processed by a capacitated system, and the *available system capacity at time t*, C_t, is defined as the maximum throughput that system resources are capable of sustaining at *t*. Consider a lean (i.e., no inventory) system with a target (demand adjusted) capacity C^* experiencing periodic random disruptions, each of which may render a full or partial system capacity loss. We assume that disruptions occur one at a time and that the i^{th} occurrence results in an instantaneous loss of magnitude ΔC_i in the remaining system capacity. Following the i^{th} disruption at time *t*, the system capacity remains at level $C_t - \Delta C_i$ until the next disruption unless the remaining capacity falls below a critical level *c* upon which the system regains all lost capacity back to C^*. For the reason of simplicity, in this paper, we assumed instantaneous recovery. The system is assumed to stochastically regenerate at points of recovery (Figure 1). Capacity dynamics as such can be observed in a number of industrial scenarios including, but are not limited to, (i) shortage of repair personnel and performance degradation caused by failing equipment with a full repair upon a complete failure, (ii) non-self-announcing stepwise system failures, and (iii) gradual equipment phaseout and modernization.

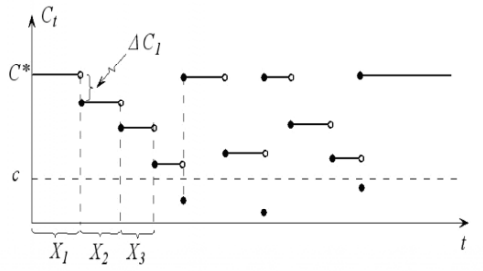


Figure 1 A realization of the system capacity dynamics

Let $\Delta C_i = \alpha_i C^*$, where $\{0 \le \alpha_i \le 1, i \in \mathbb{N}\}$ are assumed to form a sequence of i.i.d. random variables. The time of the first disruption is denoted by X_1, and $X_i, i = 2,3,\ldots$ denotes the time between $(i\text{-}1)^{th}$ and i^{th} disruptions (Figure 1).

APPENDIX C: (Continued)

We assume that $X_i, i = 2,3,\ldots$ are i.i.d. random variables. The time of the n^{th} capacity loss is expressed as $Z_n = \sum_{i=1}^{n} X_i$, $n = 1,2,...$ where we define $Z_0 = 0$. Let $N_x = \min\{n \text{ s.t.} \sum_{i=1}^{n} \Delta C_i > C^* - x\}$. It then follows that N_c is the number of capacity disruptions between two successive recovery epochs. As such, $Y = Z_{N_c}$ is the time between two successive recovery events, which marks the beginning and the end of a regenerative cycle.

A proactive decision maker has a number of mitigation options to reduce the rate of disruptions. When no countermeasures are implemented, he earns $R_t = \pi \cdot C_t$ at time t, where π is a time independent price factor minus item unit cost. Therefore, the revenue in each cycle is $R = \pi \cdot \int_0^Y C_t \mathrm{d}t = \pi \cdot C$.

We assume that a cost of $m(\lambda)$ per unit time is incurred to activate and operate a set of countermeasures that would maintain a rate of λ capacity disruptions per unit time. In this paper, we are not concerned with the description of the nature of specific countermeasure options but rather we focus on the analytics of the disruption rate reducing impact that those options have on the system performance. We assume that the decision maker has a risk-neutral utility function (Keeney & Raiffa 1993), and thus, our analysis will be based on the limiting long-run average reward as the criterion for policy assessment.

Let $\Pi = R - m(\lambda) \cdot Y$ denote the total reward earned in one renewal cycle and $\Pi_t = \int_0^t R_z \, dz - m(\lambda) \cdot t$ denote the total reward by time t . The long-run average reward converges then to the following (Ross 1996):

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)}. \tag{1}$$

In this paper, we first consider a one-phase mitigation policy in which countermeasures are activated throughout a regenerative cycle. Later, we will expand the analysis to examine a two-phase model.

3. One-Phase Countermeasure Policy

When countermeasures are engaged throughout the entire cycle, $E(\Pi) = \pi \cdot E(C) - m(\lambda) \cdot E(Y)$, and hence, we seek to derive the expected cycle length and the expected cycle capacity. We assume that interarrival times X_i are distributed exponentially with rate λ and that fractional capacity losses α_i are distributed uniformly over [0, 1]. Total capacity per cycle can be expressed as

$$C = C^* \left[\sum_{i=1}^{N_c} X_i - \sum_{i=2}^{N_c} \sum_{j=1}^{i-1} \alpha_j X_i \right], \tag{1}$$

whereas the cycle length is $Y = \sum_{i=1}^{N_c} X_i$. Before we proceed with computing $E(C)$ and $E(Y)$, we will need the following result to compute $E(N_c)$, the expected number of capacity loss events per cycle.

Result 1 *Let* $\zeta_i, i = 1,\ldots,n$ *be i.i.d. uniform* [0,1] *random variables. Then* $P(\sum_{i=1}^{n} \zeta_i \le u) = u^n / n!$

Proof. We prove by induction. For $n = 1$, the result is trivial. Assuming that the result holds for $n-1$, note that

APPENDIX C: (Continued)

$$f_{\zeta_1,\zeta_2,\ldots,\zeta_{n-1}}(u) = u^{n-2}/(n-2)!.$$

We have

$$P(\sum_{i=1}^{n}\zeta_i \le u) = P(\sum_{i=1}^{n-1}\zeta_i + \zeta_n \le u)$$

$$= \int_0^u \int_0^{u-s} dy \frac{s^{n-2}}{(n-2)!} ds$$

$$= \int_0^u (u-s)\frac{s^{n-2}}{(n-2)!} ds$$

$$= \frac{u^n}{(n-1)!} - \frac{u^n}{n(n-2)!} = \frac{u^n}{n!}. \quad \blacksquare$$

Now we are in a position to compute $E(Y)$ using Result 1. Let $c = (1-\alpha)C^*$ for some $\alpha \in (0,1)$. Note the equivalency of events $\{N_c = n\}$ and $\{\sum_{i=1}^{n-1}\alpha_i \le \alpha \text{ and } \sum_{i=1}^{n}\alpha_i > \alpha\}$.

Using Result 1 we have,

$$P(\sum_{i=1}^{n-1}\alpha_i \le \alpha \text{ and } \sum_{i=1}^{n}\alpha_i > \alpha)$$

$$= P(\sum_{i=1}^{n-1}\alpha_i \le \alpha) - P(\sum_{i=1}^{n}\alpha_i \le \alpha)$$

$$= \frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!},$$

which can be used to obtain,

$$E(N_c) = \sum_{n=1}^{\infty} n\cdot[\frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!}]$$

$$= \sum_{n=0}^{\infty}\frac{\alpha^n}{n!} = e^{\alpha}, \qquad (2)$$

and

$$E(Y) = \sum_{n=1}^{\infty} E(\sum_{i=1}^{N_c} X_i \mid N_c = n)P(N_c = n)$$

$$= \sum_{n=1}^{\infty}\frac{n}{\lambda}\cdot P(N_c = n) = \frac{1}{\lambda}\cdot E(N_c) = \frac{e^{\alpha}}{\lambda}.$$

Computation of $E(C)$ can be found in the Appendix. We have that

$$E(C) = \frac{C^*}{\lambda}\Big[\sum_{n=3}^{\infty} P(N_c = n)\cdot\Big\{n - \sum_{k=1}^{n-2} h_k(n,\alpha) - h(n,\alpha)\Big\} + P(N_c = 2)\cdot(2 - h(2,\alpha)) + (1-\alpha)\Big], \qquad (3)$$

where

$$h_k(n,\alpha) = \frac{k}{(n-k)(\frac{n}{\alpha}-1)}\Big[n - k - \alpha + \alpha\frac{k+1}{n+1}\Big],$$

and

$$h(n,\alpha) = \frac{n(n-1)}{n\alpha^{n-1} - \alpha^n}\Big[(1-\alpha)\frac{\alpha^n}{n} + \frac{\alpha^{n+1}}{n+1}\Big].$$

Using (3), we can compute the long-run average reward in the following way. Define $\bar{C}(\alpha)$ as

$$\bar{C}(\alpha) = \sum_{n=3}^{\infty} P(N_c = n)\cdot\{n - \sum_{k=1}^{n-2} h_k(n,\alpha) - h(n,\alpha)\} + P(N_c = 2)(2 - h(2,\alpha)) + (1-\alpha).$$

Then, the limiting value of long-run average reward is given by the following expression.

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)}$$

$$\frac{E(\Pi)}{E(Y)} = \frac{\frac{\pi\cdot C^*\cdot\bar{C}(\alpha)}{\lambda} - \frac{e^{\alpha}}{\lambda}\cdot m(\lambda)}{e^{\alpha}/\lambda}. \qquad (4)$$

In this section, we have considered a one-phase mitigation policy where countermeasures are implemented during the entire disruption cycle. The expression for the limiting long-run average reward (Eq. 4) will serve as a basis for analyzing a two-phase policy in the next section.

4. A Two-Phase Countermeasure Policy

Consider the set of policies under which

APPENDIX C: (Continued)

countermeasures are activated at the beginning of each system cycle and remain in effect as long as system capacity exceeds a certain higher level $c_l > c$, where $c_l = (1-\alpha_l)\cdot C^*$ for some $\alpha_l \in (0,1)$. Countermeasures remain deactivated for levels below c_l, where the system becomes exposed to "normal" disruption rate. This model is driven by the idea that from the system manager's viewpoint, it is desirable to stay longer in the "on" zone, closer to the target level C^* rather than prolong the "off" portion of the cycle. As in Section 2 and 3, c is the critical lower level that triggers instantaneous capacity recovery (Figure 2). The system is said to be "on" when countermeasures are in effect and "off" otherwise. Long-run average reward of this altered process exhibits the same convergence property. Therefore, it is our interest in this section to compute $E[\Pi]/E[Y]$.

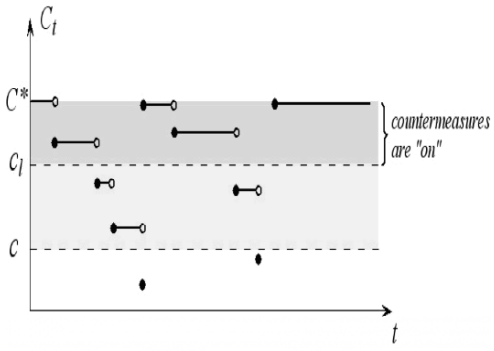


Figure 2 A realization of the system capacity dynamics for a two-phase policy. Disruption rate during the "on" phase (λ_l) is smaller than the disruption rate during the "off" phase (λ).

We first derive the distribution of the initial system capacity for the "off" period in a cycle. The following proposition summarizes the result,

Proposition 1 *Consider a capacitated system in which capacity disruption interarrival times are exponentially distributed with parameter* λ, *fractional stepwise capacity losses follow a uniform distribution on* $[0,1]$, *and capacity is restored fully and instantaneously upon falling below level* c. *Suppose that the system is "on" when* $C_t > (1-\alpha_l)\cdot C^*$, *"off" otherwise, and* $c = (1-\alpha)C^*, \alpha_l < \alpha$. *Then, the distribution of initial system capacity of the "off" period is given by the following,*

$$P\Big(\sum_{i=1}^{N_{cl}} \Delta C_i \le \bar{\alpha} C^*\Big) = (\bar{\alpha} - \alpha_l)e^{\alpha_l}.$$

Proof. We proceed by considering the number of capacity losses during the "on" period, N_{c_l}. Let $\Gamma_k = \sum_1^k \alpha_i$. For $n \ge 2$,

$$P\Big(\sum_{i=1}^{N_{cl}} \Delta C_i \le \bar{\alpha} C^* \mid N_{c_l} = n\Big)$$

$$= P(\Gamma_{n-1} \le \alpha_l, \Gamma_n \in (\alpha_l, \bar{\alpha}) \mid \Gamma_{n-1} \le \alpha_l, \Gamma_n > \alpha_l)$$

$$= P(\Gamma_{n-1} \le \alpha_l, \Gamma_n \in (\alpha_l, \bar{\alpha}))/P(\Gamma_{n-1} \le \alpha_l, \Gamma_n > \alpha_l)$$

$$= \Big[\int_0^{\alpha_l} f_{\Gamma_{n-1}}(s)\cdot P(\alpha_n \in (\alpha_l - s, \bar{\alpha} - s))ds\Big]/\Big[\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}\Big]$$

$$= (\bar{\alpha} - \alpha_l)\frac{\alpha_l^{n-1}}{(n-1)!}/\Big[\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}\Big]. \quad (5)$$

Note that

$$P(\alpha_1 \in (\alpha_l, \bar{\alpha}) \mid \alpha_1 > \alpha_l) = (\bar{\alpha} - \alpha_l)/(1-\alpha_l).$$

Therefore, one can verify by slight modifications in the computations above that (5) holds for $n = 1$ as well. Using Result 1, we

APPENDIX C: (Continued)

obtain

$$P\Big(\sum_{i=1}^{N_{c_l}} \Delta C_i \le \bar{\alpha} C^*\Big) = (\bar{\alpha} - \alpha_l) \cdot \sum_{n=1}^{\infty} P(N_{c_l} = n)$$

$$\Big\{\frac{\alpha_l^{n-1}}{(n-1)!} / \Big[\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}\Big]\Big\}$$

$$= (\bar{\alpha} - \alpha_l) \sum_{n=1}^{\infty} \frac{\alpha_l^{n-1}}{(n-1)!} = (\bar{\alpha} - \alpha_l) \cdot e^{\alpha_l},$$

which concludes the proof. ■

Expected cycle reward is the sum of expected returns of the "on" and "off" cycle periods. Let Y_l and Y_s denote the length of the "on" and "off" cycle periods, respectively, where $Y = Y_l + Y_s$ is the length of the cycle. Define

$$\overline{C_l}(\alpha_l) = \sum_{n=3}^{\infty} P(N_{c_l} = n) \cdot \Big\{n - \sum_{k=1}^{n-2} h_k(n, \alpha_l)$$

$$-h(n, \alpha_l)\Big\} + P(N_{c_l} = 2) \cdot (2 - h(2, \alpha_l)) + (1 - \alpha_l).$$

Results of the previous section can be readily applied to obtain the expression for expected total capacity, C_l, during the "on" period, which is $E(C_l) = \overline{C_l}(\alpha_l) \cdot C^* / \lambda_l$, where λ_l is disruption rate during the "on" period. Similarly, we have $E(Y_l) = e^{\alpha_l} / \lambda_l$. While the length of each cycle is affected by the change in disruption rate, the total number of disruption events in a cycle, N_c, is determined solely by a uniform capacity reduction process that evolves independently from the disruption rate. Therefore, we can deduce immediately using (2) that $E(N_c) = e^{\alpha}$ and $E(N_{c_l}) = e^{\alpha_l}$. Then, since capacity disruption rate remains at λ during the "off" period we have,

$$E(Y_s) = \frac{e^{\alpha} - e^{\alpha_l}}{\lambda}.$$

Likewise, expected total capacity during the "off" period, C_s, can be readily obtained after considering the initial capacity level in the "off" period, C_0. Note that

$$C_s = C_0 \cdot \sum_{i=1}^{N_c - N_{c_l}} X_i - C^* \cdot \sum_{i=2}^{N_c - N_{c_l}} \sum_{j=1}^{i-1} X_i \alpha_j,$$

so that

$$E(C_s \mid C_0 = (1 - \bar{\alpha}) \cdot C^*)$$

$$= C_0 \cdot \frac{E(N_c - N_{c_l} \mid C_0 = (1 - \bar{\alpha}) \cdot C^*)}{\lambda}$$

$$-\frac{C^*}{\lambda} \cdot \sum_{n=2}^{\infty} P(N_c - N_{c_l} = n \mid C_0 = (1 - \bar{\alpha}) \cdot C^*)$$

$$\cdot \sum_{i=2}^{n} E\Big(\sum_{j=1}^{i-1} \alpha_j \mid N_c - N_{c_l} = n\Big).$$

This expression can be simplified to yield the following.

$$E(C_s \mid C_0) = \frac{C_0}{\lambda} \cdot e^{\alpha - \bar{\alpha}} - \frac{C^*}{\lambda}$$

$$\Big[\sum_{n=3}^{\infty} P(N_c - N_{c_l} = n \mid C_0 = (1 - \bar{\alpha}) \cdot C^*)$$

$$\Big\{\sum_{k=1}^{n-2} h_k(n, \alpha - \bar{\alpha}) + h(n, \alpha - \bar{\alpha})\Big\} +$$

$$P(N_c - N_{c_l} = 2 \mid C_0 = (1 - \bar{\alpha}) \cdot C^*) \cdot$$

$$h(2, \alpha - \bar{\alpha})\Big]$$

$$= \frac{C_0}{\lambda} \cdot e^{\alpha - \bar{\alpha}} - \frac{C^*}{\lambda} \cdot \psi(\bar{\alpha}, \alpha).$$

We also have

$$P(N_c - N_{c_l} = n \mid C_0 = (1 - \bar{\alpha}) C^*)$$

$$= \frac{(\alpha - \bar{\alpha})^{n-1}}{(n-1)!} - \frac{(\alpha - \bar{\alpha})^n}{n!}.$$

APPENDIX C: (Continued)

In order to compute $E(C_s)$, we need the expression for $E(C_0 \cdot e^{\alpha-\bar{\alpha}})$, which is derived as follows,

$$E(C_0 \cdot e^{\alpha-\bar{\alpha}})$$

$$= E(E(C_0 \cdot e^{\alpha-\bar{\alpha}} \mid N_{c_l}))$$

$$= C^* \cdot \sum_{n=1}^{\infty} P(N_{c_l} = n) \cdot \int_{\alpha_l}^{\alpha} (1-\bar{\alpha}) \cdot e^{\alpha-\bar{\alpha}} \cdot \{\frac{\alpha_l^{n-1}}{(n-1)!} / [\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}]\} d\bar{\alpha}$$

$$= C^* \cdot (\alpha - \alpha_l \cdot e^{\alpha-\alpha_l}) \cdot \sum_{n=1}^{\infty} [\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}] \cdot \{\frac{\alpha_l^{n-1}}{(n-1)!} / [\frac{\alpha_l^{n-1}}{(n-1)!} - \frac{\alpha_l^n}{n!}]\}$$

$$= C^* \cdot (\alpha - \alpha_l \cdot e^{\alpha-\alpha_l}) \cdot e^{\alpha_l}$$

$$= C^* \cdot (\alpha e^{\alpha_l} - \alpha_l \cdot e^{\alpha}) \cdot$$

Now, we are in a position to obtain $E(C_s)$:

$$E(C_s) = \frac{C^*}{\lambda} \cdot (\alpha e^{\alpha_l} - \alpha_l e^{\alpha}) - \frac{C^*}{\lambda} \int_{\alpha_l}^{\alpha} e^{\alpha_l} \cdot \psi(\bar{\alpha}, \alpha) d\bar{\alpha}$$

$$= \frac{C^*}{\lambda} \cdot f(\alpha_l, \alpha).$$

This brings us to the following principle proposition.

Proposition 2 *For the capacitated system described in Proposition 1, the long-run average reward converges to the following expression.*

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)}$$

$$= \frac{\pi \cdot [\frac{C^* \cdot \bar{C}_l(\alpha_l)}{\lambda_l} + \frac{C^*}{\lambda} f(\alpha_l, \alpha)] - \frac{e^{\alpha_l}}{\lambda_l} \cdot m(\lambda_l)}{[e^{\alpha_l} / \lambda_l] + [(e^{\alpha} - e^{\alpha_l}) / \lambda]}. \quad (6)$$

Proof. Using Theorem 3.6.1 in (Ross 1996), we know that long-run average reward converges to,

$$\frac{\Pi_t}{t} \to \frac{\pi \cdot (E(C_l) + E(C_s)) - E(Y_l) \cdot m(\lambda_l)}{E(Y_s) + E(Y_l)}. \quad (7)$$

The proof follows by substituting expressions for $E(C_l)$, $E(C_s)$, $E(Y_l)$ and $E(Y_s)$ into (7). ■

5. Numerical Analysis and Sensitivity Study

An optimal two-phase policy maximizes long-run average reward by activating countermeasures that set optimal levels of λ_l and α_l. In what follows, we conduct a parametric analysis of the optimal policy behavior (Eq. 6). Note that in computing $f(\alpha_l, \alpha)$ (through $\psi(\bar{\alpha}, \alpha)$) and $\bar{C}_l(\alpha_l)$, we encounter infinite sums which include terms $P(N_c - N_{c_l} = n)$ and $P(N_{c_l} = n)$, respectively. These terms represent the probability distribution of the number of disruptions during the "off" and "on" phases, respectively. Since the mean disruption magnitude is strictly positive and C^* (and hence, c_l and c) is finite, both of these terms go to zero as $n \to \infty$.

Since the disrupted capacity is regained instantaneously at the end of each cycle, the solution to an optimum policy (α_l^*, α^*) is trivial. Hence, we analyze the behavior of the optimal α_l as a function of α, i.e., $\alpha_l^*(\alpha)$, for fixed values of λ and λ_l. We first observe that $\alpha_l^*(\alpha)$ is monotonically non-decreasing in α: as α increases, the cycle time will

APPENDIX C: (Continued)

increase as well, and so are the periods of lower system capacity. On the other hand, increasing α_l results in longer periods of higher system capacity. This trade-off between benefits and costs of activating countermeasures renders $\alpha_l^*(\alpha) < \alpha$. In what follows, we use the initial parameter values as shown in Table 1.

The cost of countermeasures is assumed to be of the form $m(\lambda_l) = (\lambda / \lambda_l)^r$. The cost decreases as the disruption rate λ_l gets higher, which is used to measure effectiveness of countermeasure technology, and the cost increases in r, which is used to model the marginal cost of installing a more effective technology. As Figure 3 illustrates, $\alpha_l^*(\alpha)$ is increasingly decreasing in r. We also observe that as r gets larger, $\alpha_l^*(\alpha)$ exhibits a higher sensitivity to per unit changes in r. As one can see, in flat regions of Figure 3, reducing the disruption rate is not economically sound.

Table 1 Initial parameter values

λ_l	π	λ	C^*	r
0.0005	1,000	0.001	1	1

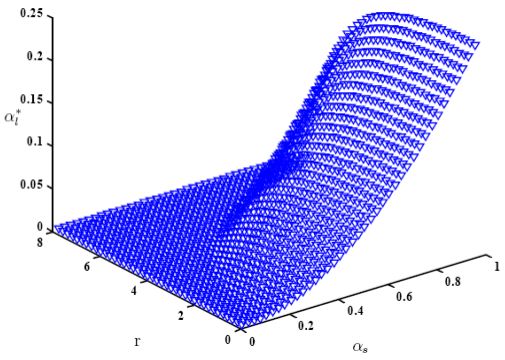


Figure 3 A 3-dimensional representation of α_l^* as a function of α and r.

For a linear cost function (r = 1, plot I in Figure 4), $\alpha_l^*(\alpha)$ is increasing in (λ_l / λ), which implies that the incremental benefits of reducing the rate of capacity disruptions do not warrant the use of countermeasures over extended periods of time. However, if the marginal cost of installing better countermeasures is not constant (r=0.001, plot II in Figure 4), the plots of $\alpha_l^*(\alpha)$ for different values of (λ_l / λ) intersect. If the marginal cost of a decreased (λ_l / λ) is relatively small, then $\alpha_l^*(\alpha)$ may be increasing with a more advanced technology (this relationship does not hold for higher values of α). However, both plots agree that the rate of increase of $\alpha_l^*(\alpha)$ is higher in the lower region of (λ_l / λ) values. Therefore, we see that expected increase in countermeasure costs over extended periods outweighs the benefits of better technology.

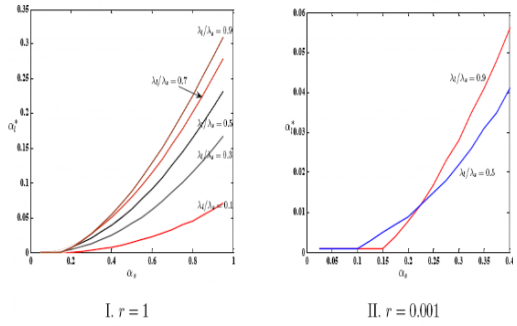
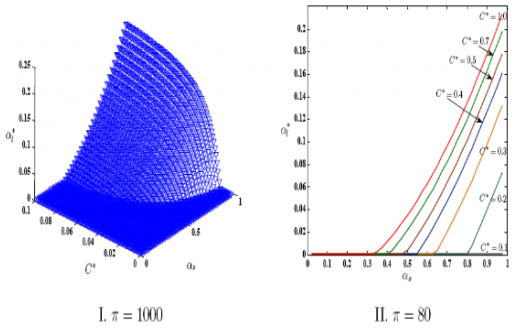


Figure 4 Behavior of $\alpha_l^*(\alpha)$ for different values of λ_l / λ.

Furthermore, we observe that $\alpha_l^*(\alpha)$ is insensitive to changes in maximum capacity C^* in the neighborhood of initial parameter values in Table 1. Common wisdom, however, suggests that C^* shall be positively correlated with the optimum period of activated countermeasures.

APPENDIX C: (Continued)

Should all items be sold, increasing C^* would lead to higher profits and hence, countermeasures should be engaged for longer periods (plot I in Figure 5, where C^* takes values in [0, 0.1]). As C^* approaches 0.1, marginal increase in $\alpha_l^*(\alpha)$ falls off sharply. This suggests that the optimal period of activated countermeasures is insensitive to changes in maximum capacity, if C^* is already high. Also, the region of sensitivity of C^* is a function of the unit profit. As illustrated in the second plot of Figure 5, $\alpha_l^*(\alpha)$ becomes responsive to changes in $C^* \in$ [0.1, 1.0], when π gets smaller (this change in sensitivity may be minimal if C^* is already high).



I. $\pi = 1000$ II. $\pi = 80$

Figure 5 Behavior of $\alpha_l^*(\alpha)$ for different values of α and C^*.

A similar relation exists between $\alpha_l^*(\alpha)$ and π. Figure 6 illustrates that C^* is insensitive to changes in π around the original parameter value of π =1000 whereas at lower unit profit levels, marginal changes in π render larger perturbations in $\alpha_l^*(\alpha)$. Changes in system capacity for low value items may require more radical changes in countermeasure policy. Nevertheless, the region of sensitivity is relatively small for both π and C^*, which suggests on a larger scale that $\alpha_l^*(\alpha)$ is quite robust to changes in system profitability.

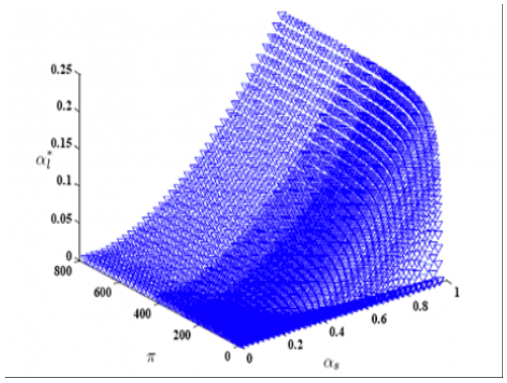


Figure 6 Behavior of $\alpha_l^*(\alpha)$ for different values of α and π.

6. Conclusions

In this manuscript, we presented one of initial attempts to fill the vacuum in the existing literature focused on development of active countermeasure policies for managing lean capacitated systems in the presence of random capacity disruptions. The system under consideration experienced stepwise partial capacity disruptions with exponentially distributed interarrival times and uniformly distributed magnitudes, followed by instantaneous recovery. Examples of such capacity dynamics include: (i) shortage of repair personnel and performance degradation caused by failing equipment with a full repair upon a complete failure, (ii) non-self-announcing stepwise system failures, and (iii) gradual equipment phaseout and modernization.

We explored two different countermeasure policies for a risk-neutral decision maker, who seeks to maximize the long-run average reward. The initial model considered a one-phase policy,

APPENDIX C: (Continued)

where countermeasures were implemented during the entirety of a disruption cycle. The results of this model served as a basis to analyze a two-phase strategy, where countermeasures were activated during only a fraction of a disruption cycle. For the latter model, we aimed to determine the optimal threshold when the countermeasures should be disengaged. In this paper, we are primarily concerned with analytics of the impact that countermeasure options can have on the system performance. In practice, the countermeasure options could range from purely technological solutions, such as installation of fire prevention water sprinkler systems, to non-technological decisions that could, for example, alleviate labor strikes or prevent terrorist attacks or political unrest. In this investigation, we considered two forms of the countermeasure cost functions. Our sensitivity analysis for the two-phase policy reveals that as the system profitability increases and the costs of countermeasures become smaller, the optimal countermeasure policy becomes less sensitive to changes in the system parameter values.

In this paper, we did not address the question of the best critical threshold that initiates immediate capacity recovery, as we assumed that the cost associated with administering any level of α is zero. Therefore, the problem of obtaining the optimal pair (α_l^*, α^*) has a trivial solution (i.e., set α^* =0). Rather, we aimed to find the optimal time in each regenerative cycle when countermeasures should be terminated given a capacity recovery threshold of α. Section 5 presented a numerical analysis to determine optimal $\alpha_l^*(\alpha)$ that maximized long-run average reward under various parametric settings. We presented the results of our sensitivity analysis for an exponential cost function. In general, $\alpha_l^*(\alpha)$ was found to be quite sensitive to exponentially increasing cost, as well as capacity and unit profit changes, if the system was already operating with low profit margins. However, as the profitability of the system increased, $\alpha_l^*(\alpha)$ had a robust response to system parameter changes.

In general, capacity disruption risk can be mitigated by reducing the probability of the hazardous events as well as their severity. In this paper, we considered countermeasures that mostly impact the probability of hazardous events rather than their severity. For risks that render partial capacity disruptions, the model recommends implementation of countermeasures during only a fraction of the operational cycle. In many cases, partial capacity disruptions are caused by risks associated with daily operations, such as small fire events and stoppages due to machine failures. For such events, our results can substantiate that certain countermeasures may be cost prohibitive even when they offer significant reduction in the disruption rate. For example, in a manufacturing facility, installation of costly fire extinguishing systems may be disfavored to employee training programs that raise awareness of overall factory cleanliness.

This paper provides one of the initial attempts for providing closed form solutions for optimal countermeasure policies for mitigation of random disruptions in capacitated systems. We hope that the presented models will be further generalized to address similar questions for capacitated systems evolving under more complex capacity dynamics. We also believe that such single-facility models will form a basis

APPENDIX C: (Continued)

to approach capacity management issues in large enterprise networks.

Acknowledgment

The authors like to thank the referees for their help to improve the quality of the paper.

Appendix

Computation of $E(C)$ in Equation 3. We begin by conditioning on N_c. Noting the dependency between ΔC_i and N_c, we first compute the conditional density of $\Gamma_k = \sum_1^k \alpha_i$, $f_{\Gamma_k}(s \mid N_c = n)$ for $k < n-1$ and $n \geq 3$. Note that

$$f_{\Gamma_k}(s \mid N_c = n) = \int_s^\alpha \frac{P(\Gamma_k = s, \Gamma_{n-1} = u, \Gamma_n > \alpha)}{P(\Gamma_{n-1} \leq \alpha, \Gamma_n > \alpha)} du. \quad (8)$$

Since α_i's are independent, we can use Result 1 to obtain

$$P(\Gamma_k = s, \Gamma_{n-1} = u, \Gamma_n > \alpha) = P(\sum_{i=1}^k \alpha_i = s) P(\sum_{i=k+1}^{n-1} \alpha_i = u) P(\alpha_n > \alpha - u)$$

$$= (1 + u - \alpha) \cdot \frac{s^{k-1}}{(k-1)!} \cdot \frac{(u-s)^{n-k-2}}{(n-k-2)!}. \quad (9)$$

Substituting (9) in (8) and evaluating the integral, we obtain

$$f_{\Gamma_k}(s \mid N_c = n) = \frac{1}{P(\Gamma_{n-1} \leq \alpha, \Gamma_n > \alpha)} \cdot \int_s^\alpha (1 + u - \alpha) \cdot \frac{s^{k-1}}{(k-1)!} \cdot \frac{(u-s)^{n-k-2}}{(n-k-2)!} du$$

$$= \frac{s^{k-1}}{(k-1)!} \cdot [(\frac{(\alpha-s)^{n-k-1}}{(n-k-1)!} - \frac{(\alpha-s)^{n-k}}{(n-k)!}) / (\frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!})].$$

We use this conditional density to compute $E(\Gamma_k \mid N_c = n)$ as follows:

$$E(\Gamma_k \mid N_c = n) = \int_0^\alpha \frac{s^{k-1}}{(k-1)!} \cdot [(\frac{(\alpha-s)^{n-k-1}}{(n-k-1)!} - \frac{(\alpha-s)^{n-k}}{(n-k)!}) / (\frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!})] \, ds.$$

$$= k \cdot \frac{n!}{k!(n-k)!} \cdot \frac{1}{n\alpha^{n-1} - \alpha^n} \cdot [(n-k-\alpha) \int_0^\alpha s^k (\alpha - s)^{n-k-1} ds + \int_0^\alpha s^{k+1} (\alpha - s)^{n-k-1} ds].$$

Note that both integrals represent beta functions (to see this, make a variable change $t = s/\alpha$). Therefore, we arrange the expressions to obtain

$$E(\Gamma_k \mid N_c = n) = k \cdot \frac{n!}{k!(n-k)!} \cdot \frac{1}{n\alpha^{n-1} - \alpha^n} \cdot [(n-k-\alpha)\alpha^n B(k+1, n-k) + \alpha^{n+1} B(k+2, n-k)]$$

$$= \frac{k}{(n-k) \cdot (\frac{n}{\alpha} - 1)} \cdot [n - k - \alpha + \alpha \frac{k+1}{n+1}]$$

$$= h_k(n, \alpha). \quad (10)$$

Note that this expression holds for $k < n-1$ and $n \geq 3$. For $k = n-1$ and $n \geq 2$,

$$f_{\Gamma_{n-1}}(s \mid N_c = n) = \frac{f_{\Gamma_{n-1}}(s) \cdot P(\alpha_n > \alpha - s)}{P(\Gamma_{n-1} \leq \alpha, \Gamma_n > \alpha)} \cdot$$

APPENDIX C: (Continued)

Using the above expression, we eventually obtain

$$E(\Gamma_{n-1} \mid N_c = n)$$

$$= \frac{1}{\frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!}} \int_0^\alpha \frac{s^{n-1}}{(n-2)!} \cdot (1-\alpha+s)\, ds$$

$$= \frac{n(n-1)}{n\alpha^{n-1} - \alpha^n} [(1-\alpha)\frac{\alpha^n}{n} + \frac{\alpha^{n+1}}{n+1}]$$

$$= h(n,\alpha). \qquad (11)$$

We can now derive expected capacity per cycle. Equation (1) gives

$$E(C) = E(C^* \cdot E(\sum_{i=1}^{N_c} X_i - \sum_{i=2}^{N_c}\sum_{j=1}^{i-1} X_j \alpha_j \mid N_c = n))$$

$$= C^* [\sum_{n=2}^{\infty} P(N_c = n) \cdot \{\frac{n}{\lambda} -$$

$$E(\sum_{i=2}^{n}\sum_{j=1}^{i-1} X_i \alpha_j \mid N_c = n)\} + \frac{P(N_c = 1)}{\lambda}]$$

$$= C^* [\sum_{n=2}^{\infty} P(N_c = n) \cdot \{\frac{n}{\lambda} -$$

$$\sum_{i=2}^{n} E(X_i) \cdot E(\Gamma_{i-1} \mid N_c = n)\} + \frac{1-\alpha}{\lambda}]$$

$$= \frac{C^*}{\lambda} [\sum_{n=2}^{\infty} P(N_c = n) \cdot \{n -$$

$$\sum_{i=2}^{n} E(\Gamma_{i-1} \mid N_c = n)\} + (1-\alpha)].$$

Finally, substituting equations (10) and (11), we obtain

$$E(C) = \frac{C^*}{\lambda} [\sum_{n=3}^{\infty} P(N_c = n) \cdot \{n -$$

$$\sum_{k=1}^{n-2} h_k(n,\alpha) - h(n,\alpha)\}$$

$$+ P(N_c = 2) \cdot (2 - h(2,\alpha)) + (1-\alpha)]$$

$$= \frac{C^*}{\lambda} E(N_c) - \frac{C^*}{\lambda} [\sum_{n=3}^{\infty} P(N_c = n) \cdot$$

$$\{\sum_{k=1}^{n-2} h_k(n,\alpha) + h(n,\alpha)\} + P(N_c = 2) \cdot h(2,\alpha)]$$

$$= \frac{C^*}{\lambda} e^{\alpha} - \frac{C^*}{\lambda} [\sum_{n=3}^{\infty} P(N_c = n) \cdot$$

$$\{\sum_{k=1}^{n-2} h_k(n,\alpha) + h(n,\alpha)\} + P(N_c = 2) \cdot h(2,\alpha)].$$

References

[1] Abboud, N.E. (1997). A simple approximation of the EMQ model with poisson machine failures. Production Planning and Control, 8: 385-397

[2] Abboud, N.E. (2001). A discrete-time Markov production-inventory model with machine breakdowns. Computers & Industrial Engineering, 39: 95-107

[3] A.M. Best Company, Inc. (2006). Managing Business Risk in 2006 and Beyond. In: Protecting Value. Available via DIALOG. http://www.protectingvalue.com. Cited September 9, 2009

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

[5] Barnes-Schuster, D., Bassok, D. & Anupindi, R. (2002). Coordination and flexibility in supply contracts with options. Manufacturing & Service Operations Management, 4: 171-207

[6] Berniker, M. (2003). SARS Morphs new concerns for Asian IT industry. In:

APPENDIX C: (Continued)

Internetnews. Available via DIALOG. http://www.internetnews.com/bus-news/article.php/2203101. Cited September 9, 2009

[7] Bielecki, R.E. & Kumar, P.R. (1988). Optimality of zero inventory policies for unreliable manufacturing systems. Operations Research, 36: 532-541

[8] Buzacott, J.A. & Shantikumar, J.G. (1993). Stochastic Models of Manufacturing Systems. Prentice Hall, Englewood Cliffs, NJ

[9] Chao, H. (1987). Inventory policy in the presence of market disruptions. Operations Research, 2: 274-281

[10] Chap, H. & Manne, S. (1982). An integrated analysis of U.S. oil stockpiling policies. Ballinder Publishing Company, Tech. rep. Cambridge, MA

[11] Cheung, K. & Lee, L. (2002). The inventory benefit of shipment coordination and stock rebalancing in a supply chain. Management Science, 48: 300-306

[12] Corbett, C.J. & DeCroix, G.A. (2001). Shared-savings contracts for indirect materials in supply chains: channel profits and environmental impacts. Management Science, 47: 881-893

[13] Groenevelt, H., Seidmann, A. & Pintelon, L. (1992). Production lot sizing with machine breakdowns. Management Science, 38: 104-123

[14] Groenevelt, H., Seidmann, A. & Pintelon, L. (1992). Production batching with machine breakdowns and safety stocks. Operations Research, 40: 959-971

[15] Henig, M. & Gerchak, Y. (1990). The structure of periodic review policies in the presence of random yield. Operations Research, 38: 634-643

[16] Hopp, W.J., Pati, N. & Jones, P.C. (1989). Optimal inventory control in a production flow system with failures. International Journal of Production Economics, 27: 1367-1384

[17] Isidore, C. (2002). Hope in West Coast port talks. In: CNN Money. Available via DIALOG. http://money.cnn.com/2002/10/02/news/economy/ports/index.htm. Cited September 9, 2009

[18] Lee, H.L. (1992). Lot sizing to reduce capacity utilization in a production process with defective items, process corrections, and rework. Management Science, 38: 1314-1328

[19] Lee, H.L., Padamanabhan, P. & Whang, S. (1997). Information distortion in supply chain: the bullwhip effect. Management Science, 43: 546-558

[20] Lemos, M. (2003). Worm exposes apathy, Microsoft flaws. In: CNET. Available via DIALOG. http://news.com.com/2100-1001-982135. Cited September 9, 2009

[21] Lemos, M. (2003). "Slammer" attacks may become way of life for Net. In: CNET. Available via DIALOG. http://news.com.com/Damage+control/2009-1001_3-983540.html. Cited September 9, 2009

[22] Keeney, R. & Raiffa, H. (1993). Decisions with Multiple Objectives: Preferences and Value Trade-offs. Cambridge University Press

APPENDIX C: (Continued)

[23] Kim, C.H. & Hong, Y. (1997). An extended EMQ model for a failure prone machine with general lifetime distribution. International Journal of Production Economics, 49: 215-223

[24] Milner, J.M. & Kouvelis, P. (2002). On the complementary value of accurate demand information and production and supplier flexibility. Manufacturing & Service Operations Management, 4: 99-113

[25] Mohebbi, E. (2003). Supply interruptions in a lost-sales inventory system with random lead time. Computers and Operations Research, 30: 824-835

[26] Murphy F.H., Toman, M.A. & Weiss, H.J. (1987). A stochastic dynamic Nash game analysis of policies for managing the strategic petroleum reserves of consuming nations. Management Science, 33 (4): 484-499

[27] Parlar, M. (1997). Continuous-review inventory problem with random supply interruptions. European Journal of Operational Research, 99: 366-385

[28] Parlar, M. & Berkin, D. (1991). Future supply uncertainty in EOQ models. Naval Research Logistics, 38: 50-55

[29] Parlar, M. & Perry, D. (1996). Inventory models of future supply uncertainty with single and multiple suppliers. Naval Research Logistics, 43: 191-210

[30] Rahim, M.A. (1994). Joint determination of production quantity, inspection schedule, and control chart design. IIE Transactions, 26: 2-11

[31] Rosenblatt, M.J. & Lee, H.L. (1986). Economic production cycles with imperfect production processes. IIE Transactions, 18: 48-55

[32] Ross, S.M. (1996). Stochastic Processes. John Wiley & Sons, Inc., New York, NY

[33] Song, J. & Zipkin, P. (1996). Inventory control with information about supply conditions. Management Science, 42: 1409-1419

[34] Teisberg, T.J. (1981). A dynamic programming model of the U.S. strategic petroleum reserve. Bell Journal of Economics, 12: 526-546

[35] Wilcox, J. (1999). Taiwan earthquake could affect holiday PC sales. In: CNET. Available via DIALOG. http://news.com.com/2100-1001-231332.html. Cited September 9, 2009

[36] Wilson, S. (2003). Political deadlock bolsters Chavez. In: Washington Post. Available via DIALOG. http://www.washingtonpost.com/ac2/wp-dyn?pagename=article&contentId=A15415-2003Jan19. Cited September 9, 2009

Niyazi Onur Bakır is a Visiting Assistant Professor at Bilkent University in Ankara, Turkey. His research interests include decision making under uncertainty, probabilistic modeling, stochastic processes, and economics of information. Prior to joining Bilkent University, he worked as a postdoctoral research associate at the University of Southern California's National Center for Risk and Economic Analysis of Terrorism Events (CREATE). In summer 2004, he worked as a visiting scholar at General Motors R&D, where he conducted research on production disruption

APPENDIX C: (Continued)

risks as a part of an enterprise risk management program. Dr. Bakır received a Ph.D. in industrial engineering and an M.S. in economics at Texas A&M University in 2004. He has taught courses at Texas A&M University, University of Southern California and Bilkent University.

Alex Savachkin is an Assistant Professor of the Department of Industrial & Management Systems Engineering at the University of South Florida, Tampa. His research interests include analytical support of enterprise risk analysis, public health disaster mitigation and emergency planning, and healthcare engineering. Dr. Savachkin received his Ph.D. in industrial engineering at Texas A&M University in 2005.

Andrés Uribe-Sánchez is a doctoral student of the Department of Industrial & Management Systems Engineering at the University of South Florida, Tampa. His research interests include risk-based analysis of random capacity dynamics in capacitated enterprise networks and development of dynamic mitigation strategies for cross-regional pandemic outbreaks.

APPENDIX D:
PUBLICATION 3: AN OPTIMAL COUNTERMEASURE POLICY TO MITIGATE RANDOM CAPACITY DISRUPTIONS IN A PRODUCTION SYSTEM

In this appendix, we present the final version of the manuscript "An Optimal Countermeasure Policy to Mitigate Random Capacity Disruptions in a Production System" published in the International Journal of Agile Systems and Management, volume 3, number 1/2, pages 4-17-226, 2008, by Inderscience Publisher. The co-authors, Dr. Alex Savachkin and Dr. Niyazi Bakir, authorized to include this document in my dissertation. Inderscience Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in my Ph.D. dissertation is attached in Appendix A.

APPENDIX D (Continued)

An optimal countermeasure policy to mitigate random capacity disruptions in a production system

Alex A. Savachkin*

Department of Industrial and
Management Systems Engineering,
University of South Florida,
4202 E, Fowler Avenue, ENB118,
Tampa, FL 33620-5350, USA
Fax: +1-813-974-5953
E-mail: savachki@eng.usf.edu
*Corresponding author

Niyazi Onur Bakir

Center for Homeland Security,
University of Southern California,
3710 Mc Clintock Ave, Room 322,
Los Angeles, CA 90089-2902, USA
E-mail: nbakir@usc.edu

Andres Uribe-Sanchez

Department of Industrial and
Management Systems Engineering,
University of South Florida,
4202 E. Fowler Avenue, ENB118,
Tampa, FL 33620-5350, USA
Fax: +1-813-974-5953
E-mail: auribe@mail.usf.edu

Abstract: In this paper, we investigate a manufacturing system exposed to unpredicted capacity disruptions with exponentially distributed interoccurrence times and uniformly distributed magnitudes of disruptions. Each disruption renders a stepwise partial system capacity loss accumulating over time until the remaining capacity reaches a certain level, upon which the system gradually restores the lost capacity to the target level. We examine implementation of a countermeasure policy, aimed at reducing the disruption rate, for a risk-neutral decision maker who seeks to maximise long-run average return. We explore how the policy of maintaining the optimal disruption rate is affected by a number of system parameters.

Keywords: production system; capacity; random disruptions; stepwise loss; linear recovery; convex cost function; countermeasure policy; mitigation strategy.

APPENDIX D (Continued)

Reference to this paper should be made as follows: Savachkin, A.A., Bakir, N.O. and Uribe-Sanchez, A. (2008) 'An optimal countermeasure policy to mitigate random capacity distruptions in a production system', *Int. J. Agile Systems and Management*, Vol. 3, Nos. 1/2, pp.4–17.

Biographical notes: Alex A. Savachkin is an Assistant Professor in the Department of Industrial and Management Systems Engineering, University of South Florida at Tampa. He received BS in Industrial Engineering from Belarussian State University in 1996, MS in Management of Technology from the University of Colorado in 1998, and PhD in Industrial Engineering from Texas A&M University in 2005. His research interests focus on analytical and computational support of enterprise risk management, supply chain design and disaster management. He has published in IIE Transactions.

Niyazi Onur Bakir is a Postdoctoral Research Associate at the University of Southern California Center for Homeland Security (CREATE). His area of study is economics of border security and evaluation of alternatives to prevent illegal introduction of weapons into United States. His research interests include decision making under uncertainty, probabilistic modelling, stochastic processes and economics of information. He received BS in Industrial Engineering from Bilkent University in Turkey in 1998, MS in Economics from Texas A&M University in 2000 and a PhD in Industrial Engineering from Texas A&M University in 2004.

Andres Uribe-Sanchez is a doctoral student in the Department of Industrial and Management Systems Engineering, University of South Florida at Tampa. He received BS in Industrial Engineering from the University of Los Andes, Colombia in 2003, and MS in Management Systems from the University of Puerto Rico in 2006.

---

1 Introduction

In 1995, an earthquake hit the port town of Kobe, Japan, razed to the ground 100,000 buildings and shut down Japan's largest port for over two years. In September 2002, longshoremen on the US West Coast were locked out in a labour strike for 11 days, forcing the shutdown of 29 ports. With more than $300 billion of goods shipped annually through these ports, the dispute caused between $11 and $22 billion in lost sales, spoiled perishables and underutilised capacity. The recent 2003 outbreak of SARS in China and Singapore forced Motorola to close several plants (Berniker, 2003). Man-made disasters are also on the rise, from terrorist attacks to computer viruses (Lemos 2003a,b). As evidenced from these and other examples, enterprises are consolidating their internal and external suppliers to gain economies of scale at the expense of heightened risk exposure to supply chain disruptions.

Computational models and methodologies suitable for analysing the effects of capacity imbalance have been largely unavailable (Savachkin and Wortman, 2006). While the supply chain literature abounds, only a small portion of it has been dedicated to modelling the impact of various disruptions such as demand patterns, supplier and production lead times, prices, imperfect process quality and process yield. The research efforts have been focused primarily on modelling operational disruptions and localised issues of lot sizing, scheduling, ordering policy, management of costs, etc. The work on

APPENDIX D (Continued)

modelling local supply rate changes were pioneered by Meyer et al. (1979), and later extended by Posner and Berg (1989) and Arreola-Risa and DeCroix (1998). Parlar and Berkin (1991) studied the classic EOQ problem with supply disruptions, later extended by Parlar and Perry (1996). Other work on production-inventory systems with deterministic demand and supply disruptions for a localised entity includes Bielecki and Kumar (1988), Weiss and Rosenthal (1992) and Mohebbi (2003). Papers addressing both supply disruptions and random demand include Chao (1987), Parlar (1997) and Song and Zipkin (1996).

Modelling production rate disruptions, such as machine failures, challenged many researchers for several decades, and numerous research efforts have been devoted to extending classical Economic Manufacturing Quantity (EMQ) model (Abboud, 1997; Buzacott and Shanthikumar, 1993; Groenevelt et al., 1992; Henig and Gerchak, 1990; Hopp et al., 1989; Kim and Hong, 1997; Lee, 1992; Porteus, 1986; Rosenblatt and Lee, 1986). As an example, Gallego (1988a,b) examine the classical economic lot-sizing model with single and multiple disruptions. A single localised unreliable bottleneck facility with a constant production and demand rate that is subject to random disruptions was analysed in Moinzadeh and Aggarwal (1997). Temporary price changes have also attracted interest among operations researchers (Arcelus and Srinivasan, 1995; Ardalan, 1995; Aull-Hyde, 1992; Taylor and Bradley, 1985; Tersine and Barman, 1995).

Our literature survey reveals that research efforts addressing the disruption of supply are insufficient. Most of the open literature focuses on reactive management of inventory, ordering policies, production lot sizing, scheduling and supply chain design. To the best of our knowledge, there have been no attempts to consider introduction of active countermeasure policies to mitigate unpredicted capacity disruptions in a production system. In this paper, we offer a modelling paradigm suitable for capturing a certain regenerative stochastic trajectory of capacity in a manufacturing facility exposed to hazardous events. The mitigation strategy focuses on decreasing the capacity disruption rate. The per unit time cost of countermeasures is modelled as a decreasing convex function of the rate. The rest of the paper has the following organisation. In Section 2, we present capacity-related terminology and notation, and state model assumptions. In Sections 3 and 4, we derive the average long-run reward as a function of the capacity disruption rate. In Section 5, we analyse how the optimal disruption rate behaves as we vary

1 the unit profit

2 the maximum capacity level

3 the recovery rate

4 the rate of recovery delay.

Section 6 offers our concluding remarks.

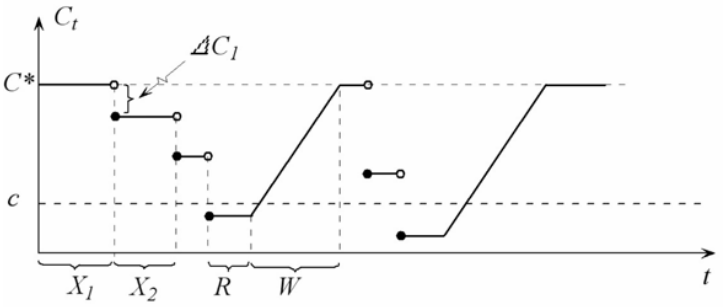
2 Problem

We consider a manufacturing system and focus on its stochastic capacity, trajectory shaped in particular, by disruptions of random origin and mode. Let C_t denote *available system capacity* at time $t \geq 0$, which is defined as the maximal throughput that system is internally capable of sustaining at t. Throughput is defined as usual as the long-run

APPENDIX D (Continued)

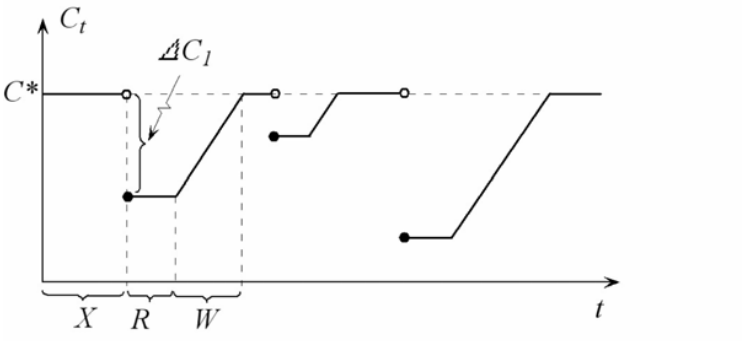
average of the number of item units per unit time processed by the system. The system operates at a target capacity C^* and experiences some periodic and unpredicted disruptions, each of which renders an instantaneous partial or full loss in system capacity. Disruptions are assumed to arrive one at a time, where the ith instance results in a decrease ΔC_i in the remaining system capacity. Capacity stays at the reduced level until the next arrival unless the remaining disrupted capacity falls below a critical threshold $c > 0$, upon which the system experiences some recovery delay followed by a gradual restoration of accumulated lost capacity back to the level C^* (Figure. 1). We assume that the system stochastically regenerates at the points of full recovery where each such point marks the beginning of a new regeneration cycle.

Figure 1 A realisation of the system capacity trajectory



The capacity trajectory described above subsumes as a special case the dynamics presented in Figure 2 (with $c = C^*$), where *every* degradation entails a recovery delay and a subsequent capacity comeback. Capacity dynamics as in Figure 2 can be observed in a number of manufacturing scenarios. An instantaneous capacity loss can be attributed to disruptions in supporting infrastructure (e.g. power outages, intermittent compressed air supply, fires, etc.), labour strikes, and equipment failure or modernisation followed by a gradual capacity recovery. Another example is a warning of act of terrorism that ensues checking the area and a piecemeal unfolding of human and manufacturing resources. In what follows, we first focus on the general case (Figure 1) and introduce modelling assumptions on capacity kinesics and mitigation options.

Figure 2 A realisation of the system capacity trajectory ($c = C^*$)



APPENDIX D (Continued)

Consider the first cycle, and let $\Delta C_i = \alpha_i C^*$, where $\{0 \le \alpha_i \le 1, i \in N\}$ are i.i.d. random variables uniformly distributed over [0,1]. Let X_1 be the time epoch of the first disruption, and let X_i, $i \ge 2$ be an interarrival time between $(i-1)$th and ith disruptions. We assume that X_i, $i \ge 1$ are i.i.d. random variables distributed exponentially with a rate $\lambda > 0$. The time of the nth capacity loss in each cycle is denoted by $Z_n = \sum_{i=1}^{n} X_i$, where $Z_0 = 0$. Let R represent a recovery delay distributed exponentially with a rate $\theta > 0$. Finally, we assume a constant rate $\gamma > 0$ of capacity recovery with W designated as the recovery period (Figure 1). The rates λ, θ, and γ are time-independent.

Let $N_x = \min\left\{\text{n s.t. } \sum_{i=1}^{n} \Delta C_i \ge \text{C*} - x\right\}$. Then, N_c is the number of disruptions during the cycle. We have that

$$W = \frac{1}{\gamma}\sum_{i=1}^{N_c} \Delta C_i$$

and hence the cycle time is $Y = Z_{N_c} + R + W$.

We assume that a proactive system manager can integrate certain countermeasure options to maintain the rate of capacity disruptions at a fixed level λ, at a cost of $c(\lambda)$ per unit time. If λ' denotes a 'regular' disruption rate (with no countermeasure options available), we assume that $c(\lambda) = 0$ for $\lambda \ge \lambda'$. We further assume that $c(\cdot)$ is convex and decreasing in λ. The mitigation strategy assumes that the countermeasures are deactivated during the recovery delay and recovery period. When the countermeasures are deactivated, the manager earns a profit πC_t at time t, where π is a time independent unit price factor minus unit cost. Thus, the total reward earned during a renewal cycle is $\Pi = \pi \int_0^Y C_t dt - c(\lambda) Z_{N_c} = \pi C - c(\lambda) Z_{N_c}$. The system manager's *objective* is to maximise the long-run average reward, where the decision variable is λ. Towards this end, using the result of Theorem 3.6.1 in Ross (1996), we observe that the long-run average reward converges to

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)}$$

Hence, in the next section, we aim to derive the expected cycle length and the expected cycle capacity.

3 Long-run average cycle capacity

The total capacity per cycle can be expressed as

$$C = C_{Z_{N_c}} + C_R + C_W$$

where $C_{Z_{N_c}}$, C_R, and C_W represent total capacity associated with Z_{N_c}, R and W portions of the cycle period, respectively. In the Appendix, we calculate $E(C_{Z_{N_c}}) = \bar{C}(\alpha)\ C^*/\lambda$, where $\bar{C}(\alpha)$ is defined accordingly as well. Next, we obtain $E(C_R)$ as follows.

APPENDIX D (Continued)

$$E(C_R) = E\left(\left[C^* - \sum_{i=1}^{N_c}\Delta C_i\right] R\right)$$
$$= \mathrm{E}\left(\left[C^* - \sum_{i=1}^{N_c}\alpha_i C^*\right] E(R)\right)$$
$$= \frac{C^* - C^* E\left(\sum_{i=1}^{N_c}\alpha_i\right)}{\theta}$$

where

$$E\left(\sum_{i=1}^{N_c}\alpha_i\right) = \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{N_c}\alpha_i \mid N_c = n\right) P(N_c = n)$$
$$= \sum_{n=1}^{\infty}\frac{n}{2} P\left(N_c = n\right)$$
$$= \frac{1}{2}E\left(N_c\right) = \frac{e^{\alpha}}{2}$$

where $E(N_c) = e^{\alpha}$, follows from (6) in Appendix. Therefore

$$E\left(C_R\right) = \frac{C^*}{\theta}\left(1 - \frac{e^{\alpha}}{2}\right)$$

Next, we calculate $E(C_W)$. To this end, we first obtain C_W as follows:

$$C_W = \left(C^* - \sum_{i=1}^{N_c}\alpha_i C^*\right) W + \frac{\gamma W^2}{2}$$
$$= \left(C^* - \sum_{i=1}^{N_c}\alpha_i C^*\right) \frac{\sum_{i=1}^{N_c}\alpha_i C^*}{\gamma} + \frac{\gamma\left[\sum_{i=1}^{N_c}\alpha_i C^*/\gamma\right]^2}{2}$$
$$= \frac{C^{*2}\sum_{i=1}^{N_c}\alpha_i}{\gamma} - \frac{C^{*2}\left[\sum_{i=1}^{N_c}\alpha_i\right]^2}{\gamma} + \frac{C^{*2}\left[\sum_{i=1}^{N_c}\alpha_i\right]^2}{2\gamma}$$
$$= \frac{C^{*2}\sum_{i=1}^{N_c}\alpha_i}{\gamma} - \frac{C^{*2}\left[\sum_{i=1}^{N_c}\alpha_i\right]^2}{2\gamma}.$$

It follows then that

$$E(C_W) = \frac{C^{*2}}{2\gamma}e^{\alpha} - \frac{C^*}{2\gamma}E\left(\left[\sum_{i=1}^{N_c}\alpha_i\right]^2\right)$$
$$= \frac{C^{*2}}{2\gamma}e^{\alpha} - \frac{C^*}{2\gamma}\left(\frac{e^{2\alpha}}{4} + \mathrm{Var}\left(\sum_{i=1}^{N_c}\alpha_i\right)\right)$$

$$= \frac{C^{*2}}{2\gamma} e^{\alpha} - \frac{C^{*}}{2\gamma}\left(\frac{e^{2\alpha}}{4} + E\left(\mathrm{Var}\left(\sum_{i=1}^{N_c} \alpha_i \mid N_c = n\right)\right)\right)$$

$$= \frac{C^{*2}}{2\gamma} e^{\alpha} - \frac{C^{*}}{2\gamma}\left(\frac{e^{2\alpha}}{4} + \frac{e^{\alpha}}{12}\right)$$

Therefore,

$$E(C) = \frac{\bar{C}(\alpha)C^{*}}{\lambda} + \frac{C^{*}}{\theta}\left(1 - \frac{e^{\alpha}}{2}\right) + \frac{C^{*2}}{2\gamma} e^{\alpha} - \frac{C^{*}}{2\gamma}\left(\frac{e^{2\alpha}}{4} + \frac{e^{\alpha}}{12}\right)$$

In the next section, we derive the expected cycle length and finalise computation of the long-run average reward limit.

4 Long-run average reward

We calculate $E(W)$ as

$$E(W) = \frac{1}{\gamma} E\left(\sum_{i=1}^{N_c} \alpha_i C^{*}\right)$$

$$= \frac{C^{*}}{\gamma} \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{N_c} \alpha_i \mid N_c = n\right) P\left(N_c = n\right)$$

$$= \frac{C^{*}}{2\gamma} E\left(N_c\right) = \frac{e^{\alpha} C^{*}}{2\gamma}$$

where $E(N_c) = e^{\alpha}$, follows from (6) in Appendix. We then have that

$$E(Y) = E\left(Z_{N_c}\right) + E\left(R\right) + E\left(W\right)$$

$$= \frac{e^{\alpha}}{\lambda} + \frac{1}{\theta} + \frac{e^{\alpha} C^{*}}{2\gamma}$$

where $E\left(Z_{N_c}\right) = e^{\alpha}/\lambda$ follows from (7) in Appendix.

At this point, we are in a position to analyse the limiting behaviour of the long-run average reward which is given by the following expression

$$\frac{\Pi_t}{t} \rightarrow \frac{E(\Pi)}{E(Y)}$$

$$= \frac{\pi\left(\left(\bar{C}(\alpha)C^{*}/\lambda\right) + \left(C^{*}/\theta\right)\left[1 - \left(e^{\alpha}/2\right)\right] + \left(C^{*2} e^{\alpha}/2\gamma\right) - \left(C^{*}/2\gamma\right)\left(\left(e^{2\alpha}/4\right) + \left(e^{\alpha}/12\right)\right)\right) - \left(e^{\alpha}/\lambda\right)c(\lambda)}{\left(e^{\alpha}/\lambda\right) + \left(1/\theta\right) + \left(e^{\alpha}C^{*}/2\gamma\right)}$$

To simplify the expression and gain meaningful insights, we limit our attention to the special case ($c = C^{*}$) as presented in Figure 2. Then the long-run average reward simplifies to

APPENDIX D (Continued)

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)} = \frac{\pi C^* \left((1/\lambda)+(1/2\theta)+\left(C^*/2\gamma\right)-(1/6\gamma)\right)-(c(\lambda)/\lambda)}{(1/\lambda)+(1/\theta)+\left(C^*/2\gamma\right)}$$

or multiplying both the numerator and the denominator by λ,

$$\frac{\Pi_t}{t} \to \frac{E(\Pi)}{E(Y)} = \frac{\left[\lambda\pi C^* \left((1/2\theta)+\left(C^*/2\gamma\right)-(1/6\gamma)\right)-c(\lambda)+\pi C^*\right]}{\left[1+\lambda\left((1/\theta)+\left(C^*/2\gamma\right)\right)\right]} \tag{1}$$

To find the maximum long-run average reward, we equalise the first derivative of (1) with respect to λ, to zero. After simplification, we obtain

$$\pi C^* \left(\frac{1}{6\gamma}+\frac{1}{2\theta}\right) = \left(\frac{1}{\theta}+\frac{C^*}{2\gamma}\right)\left(c(\lambda)-\lambda c'(\lambda)\right)-c'(\lambda) \tag{2}$$

The assumptions we made on $c(\cdot)$ render both sides of (2) positive. To see this, one needs to take the Taylor expansion to estimate $c(\lambda') = 0$ using λ as the reference point, and observe that $c(\lambda) - \lambda \cdot c'(\lambda)$ is positive since $c'(\cdot) < 0$. In the next section, we will analyse the results of comparative statics based on this equation to see how λ behaves as a function of

1 the unit profit π

2 the maximum capacity level C^*

3 the recovery rate γ

4 the rate of recovery delay θ.

5 Behaviour of the optimal disruption rate

As stated previously, the cost function $c(\lambda)$ has two intervals. For $\lambda \le \lambda'$, $c(\lambda)$ is convex and decreasing in λ (with $c(\lambda') = 0$), whereas $c(\lambda) = 0$ for $\lambda > \lambda'$. In this section, we assume that a solution λ^* to (2) exists and satisfies $\lambda^* < \lambda'$. Note that λ appears only on the right hand side of (2) (hereafter denoted as RHS). Since $c(\cdot)$ is convex, we have that

$$\frac{\partial \text{RHS}}{\partial \lambda} = -\left(\frac{1}{\theta}+\frac{C^*}{2\gamma}\right)\lambda c''(\lambda) - c''(\lambda) < 0 \tag{3}$$

Note that (3) also ensures that the solution λ^* to (2), if it exists, is unique. We first consider the impact of the unit profit π, which appears only on the left hand side of (2) (hereafter denoted as LHS). Note that LHS is increasing in π. Thus, as π increases, λ has to decrease to satisfy (2). In other words, as the unit profit increases, disruptions become more costly, and hence λ^* becomes smaller to reduce the number of disruptions over a long time horizon.

Intuitively, increasing the maximum system capacity C^* should have a similar impact on λ^*. As the system profitability increases, λ should be reduced to ensure that the system operates at the maximum capacity level. This intuitive relationship follows

APPENDIX D (Continued)

mathematically only for a range of parameter values. Note that both RHS and LHS are increasing in C^*. We have that

$$\frac{\partial \text{LHS}}{\partial C^*} = \pi\left(\frac{1}{6\gamma}+\frac{1}{2\theta}\right) > 0, \quad \frac{\partial \text{RHS}}{\partial C^*} = \frac{c(\lambda)-\lambda c'(\lambda)}{2\gamma} > 0 \tag{4}$$

Note that both derivatives in (4) are independent of C^*. If $\partial\text{LHS}/\partial C^* > \partial\text{RHS}/\partial C^*$, then LHS exceeds RHS as C^* is increased from the initial level. Therefore, the new λ^* has to be smaller. This implies that if the system profitability is sufficiently high, λ^* is a decreasing function of C^*. Furthermore, if the ratio of γ to θ exceeds a certain threshold, the same relationship holds between λ^* and C^*. If the expected recovery delay is short, the loss induced by a single disruption is not high enough to justify a lower λ^* when C^* is increased.

An increased rate of linear recovery γ clearly benefits the system performance. However, note that γ appears in both RHS and LHS of (2). Taking the first order derivatives with respect to γ yields

$$\frac{\partial \text{LHS}}{\partial \gamma} = -\pi\, \text{C}^* \frac{1}{6\gamma^2} < 0, \; \frac{\partial \text{RHS}}{\partial \gamma} = -\,\text{C}^* \frac{c(\lambda)-\lambda c'(\lambda)}{2\gamma^2} < 0$$

It follows then that $\partial\text{LHS}/\partial\gamma < \partial\text{RHS}/\partial\gamma$ for all values of γ, if $\pi > 3\,(c(\lambda) - \lambda\ c'(\lambda)\,)$. This implies that if the unit profit is sufficiently high, then λ^* is increasing in γ. When the linear recovery rate is high, the loss from a disruption event is relatively low and hence are reduced countermeasure benefits. Therefore, the system manager shall not be motivated to maintain a low level of λ^* as the economic benefits no longer outweigh the costs of such a policy.

Finally, we discuss the relationship between λ^* and θ. Similar to γ, the parameter θ appears in both RHS and LHS. We have that

$$\frac{\partial \text{LHS}}{\partial \theta} = -\pi\, \text{C}^* \frac{1}{2\theta^2} < 0, \; \frac{\partial \text{RHS}}{\partial \theta} = -\frac{c(\lambda)-\lambda c'(\lambda)}{\theta^2} < 0$$

Since $\partial\text{LHS}/\partial\theta < \partial\text{RHS}/\partial\theta$ follows independent of θ for $\pi\, C^* > 2\cdot(c\,(\lambda) - \lambda\cdot c'\,(\lambda))$, we conclude that higher values of θ render a higher λ^*, when system profitability is high. The intuition behind this relationship is simple. Higher values of θ mean reduced expected recovery delays, which in turn reduce the economic benefits of implementing countermeasures. Hence, it is economically sound to choose a higher λ^*.

6 Conclusions

In this paper, we examined a production system experiencing periodic capacity disruptions, each of which is followed by a random recovery delay and a constant linear rate of recovery. The system manager's objective was to implement a countermeasure strategy to alleviate the rate of disruptions λ with a convex cost function that is decreasing in λ. We derived an optimal level of λ that maximises the long-run average reward. The results of comparative statics suggest that choosing to maintain a lower disruption rate is optimal, if the system profitability is high. We concluded that higher unit profits and maximum capacity levels increase the costs of disruptions and hence,

APPENDIX D (Continued)

must be balanced by appropriate countermeasure strategies. Since shorter expected recovery delay and faster linear recovery reduce the economic loss of disruptions, the optimal level of λ is increasing in both parameters.

Acknowledgement

This work was supported in part by the National Science Foundation Grant # DMI 0621030.

References

Abboud, N.E. (1997) 'A simple approximation of the EMQ model with poisson machine failures', *Production Planning and Control*, Vol. 8, pp.385–397.

Arcelus, F.J. and Srinivasan, G. (1995) 'Discount strategies for one-time only sales', *IIE Transactions*, Vol. 27, pp.618–624.

Ardalan, A. (1995) 'A comparative analysis of approaches for determining optimal price and order quantity when a sale increases demand', *European Journal of Operational Research*, Vol. 84, pp.416–430.

Arreola-Risa, A. and DeCroix, C.A. (1998) 'Inventory management under random supply disruptions and partial backorders', *Naval Research Logistics*, Vol. 45, pp.687–703.

Aull-Hyde, R.L. (1992) 'Evaluation of supplier-restricted purchasing options under temporary price discounts', *IIE Transactions*, Vol. 24, pp.184–186.

Berniker, M. (2003) *SARS Morphs New Concerns for Asian IT Industry*, Available at: http://www.internetnews.com/bus-news/article.php/2203101.

Bielecki, R.E. and Kumar, P.R. (1988) 'Optimality of zero inventory policies for unreliable manufacturing systems', *Operations Research*, Vol. 36, pp.532–541.

Buzacott, J.A. and Shanthikumar, J.G. (1993) *Stochastic Models of Manufacturing Systems*, Englewood Cliffs, NJ: Prentice Hall.

Chao, H. (1987) 'Inventory policy in the presence of market disruptions', *Operations Research*, Vol. 2, pp.274–281.

Gallego, G. (1988a) 'Linear control policies for scheduling a single facility after an initial disruption', *Technical Report No. 770*, School of OR and IE, Cornell University, Ithaca, NY.

Gallego, G. (1988b) 'Produce-up-to policies for scheduling a single facility after an initial disruption', *Technical Report No. 771*, School of OR and IE, Cornell University, Ithaca, NY.

Groenevelt, H., Seidmann, A. and Pintelon, L. (1992) 'Production lotsizing with machine breakdowns', *Management Science*, Vol. 38, pp.104–123.

Henig, M. and Gerchak, Y. (1990) 'The structure of periodic review policies in the presence of random yield', *Operations Research*, Vol. 38, pp.634–643.

Hopp, W.J., Pati, N. and Jones, P.C. (1989) 'Optimal inventory control in a production flow system with failures', *International Journal of Production Research*, Vol. 27, pp.1367–1384.

Kim, C.H. and Hong, Y. (1997) 'An extended EMQ model for a failure prone machine with general lifetime distribution', *International Journal of Production Economics*, Vol. 49, pp.215–223.

Lee, H.L. (1992) 'Lot sizing to reduce capacity utilization in a production process with defective items, process corrections, and rework', *Management Science*, Vol. 38, pp.1314–1328.

Lemos, R. (2003a) *'Slammer' Attacks may become Way of Life for Net*, Available at: http://news.com.com/Damage+control/2009-1001_3-983540.html.

Lemos, R. (2003b) *Worm Exposes Apathy, Microsoft Flaws*, Available at: http://news.com.com/2100-1001-982135.html.

APPENDIX D (Continued)

Meyer, R.R., Rothkpof, M.H. and Smith, S.A. (1979) 'Reliability and inventory in a production storage system', *Management Science*, Vol. 25, pp.799–807.

Mohebbi, E. (2003) 'Supply interruptions in a lost-sales inventory system with random lead time', *Computers and Operations Research*, Vol. 30, pp.824–835.

Moinzadeh, K. and Aggarwal, P. (1997) 'Analysis of a production/inventory system subject to random disruptions', *Management Science*, Vol. 43, pp.1577–1588.

Parlar, M. (1997) 'Continuous-review inventory problem with random supply interruptions', *European Journal of Operational Research*, Vol. 99, pp.366–385.

Parlar, M. and Berkin, D. (1991) 'Future supply uncertainty in EOQ models', *Naval Research Logistics*, Vol. 38, pp.50–55.

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

Porteus, E.L. (1986) 'Optimal lot sizing, process quality improvement and setup cost reduction', *Operations Research*, Vol. 34, pp.137–144.

Posner, M.J. and Berg, M. (1989) 'Analysis of a production-inventory system with unreliable production facility', *Operations Research Letters*, Vol. 8, pp.339–345.

Rosenblatt, M.J. and Lee, H.L. (1986) 'Economic production cycles with imperfect production processes', *IIE Transactions*, Vol. 18, pp.48–55.

Ross, S.M. (1996) *Stochastic Processes*, New York, NY: John Wiley & Sons, Inc.

Savachkin, A. and Wortman, M. (2006) 'Capacity disruption in supply networks: a concise literature survey', *Under Review in Production and Operations Management*, Available at: http://www.ieedge.com/papers/survey.pdf.

Song, J. and Zipkin, P. (1996) 'Inventory control with information about supply conditions', *Management Science*, Vol. 42, pp.1409–1419.

Taylor, S.G. and Bradley, C.E. (1985) 'Optimal ordering strategies for announced price increases', *Operations Research*, Vol. 33, pp.312–325.

Tersine, R.J. and Barman, S. (1995) 'Economic purchasing strategies for temporary price discounts', *European Journal of Operational Research*, Vol. 80, pp.328–343.

Weiss, H.J. and Rosenthal, E.C. (1992) 'Optimal ordering policies when anticipating a disruption in supply or demand', *European Journal of Operational Research*, Vol. 59, pp.370–382.

APPENDIX D (Continued)

Appendix

Computation of $E(C_{Z_{N_c}})$.

We begin by noting that $C_{Z_{N_c}}$ can be expressed as

$$C_{Z_{N_c}} = C^* \left[\sum_{i=1}^{N_c} X_i - \sum_{i=2}^{N_c} \sum_{j=1}^{i-1} \alpha_j X_i \right] \tag{A1}$$

We first need the following result

Result 1: *Let* ζ_i, $i = 1,\ldots,n$ *be i.i.d. random variables where* $\zeta_i \sim [0,1]$. *Then* $P(\sum_{i=1}^{n} \zeta_i \le u) = u^n / n!$

Proof: We prove by induction. For $n = 1$, the result is trivial. Assuming that the result holds for $n - 1$, note that $f_{\zeta_1,\zeta_2,\ldots,\zeta_{n-1}}(u) = u^{n-2}/(n-2)!$ We have

$$P\left(\sum_{i=1}^{n} \zeta_i \le u\right) = P\left(\sum_{i=1}^{n-1} \zeta_i + \zeta_n \le u\right) = \int_0^u \int_0^{u-s} \mathrm{d}y \frac{s^{n-2}}{(n-2)!} \mathrm{d}s$$
$$= \int_0^u (u-s) \frac{s^{n-2}}{(n-2)!} \mathrm{d}s = \frac{u^n}{(n-1)!} - \frac{u^n}{n(n-2)!} = \frac{u^n}{n!} \quad \square$$

Now we are in a position to compute $E(Z_{N_c})$ using Result 1. Let $c = (1 - \alpha)C^*$ for some $\alpha \in (0,1)$. Note, the equivalency of events $\{N_c = n\}$ and $\{\sum_{i=1}^{n-1} \alpha_i \le \alpha \text{ and } \sum_{i=1}^{n} \alpha_i > \alpha\}$. Using Result 1 we have

$$P\left(\sum_{i=1}^{n-1} \alpha_i \le \alpha \text{ and } \sum_{i=1}^{n} \alpha_i > \alpha\right) = P\left(\sum_{i=1}^{n-1} \alpha_i \le \alpha\right) - P\left(\sum_{i=1}^{n} \alpha_i \le \alpha\right) = \frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!}$$

which can be used to obtain $E(N_c)$, the expected number of capacity loss events per cycle

$$E(N_c) = \sum_{n=1}^{\infty} n \left[\frac{\alpha^{n-1}}{(n-1)!} - \frac{\alpha^n}{n!} \right] = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!} = e^a \tag{A2}$$

and, therefore

$$E(Z_{N_c}) = \sum_{n=1}^{\infty} E\left(\sum_{i=1}^{N_c} X_i \mid N_c = n \right) P(N_c = n)$$
$$= \sum_{n=1}^{\infty} \frac{n}{\lambda} P(N_c = n) = \frac{1}{\lambda} E(N_c) = \frac{e^\alpha}{\lambda} \tag{A3}$$

We proceed by conditioning on N_c. Noting the dependency between ΔC_i and N_c, we first compute the conditional density of $\Gamma_k = \sum_1^k \alpha_i$, $f_{\Gamma_k}(s \mid N_c = n)$ for $k < n - 1$ and $n \ge 3$.

APPENDIX D (Continued)

Note that

$$
\begin{aligned}
f_{\Gamma_k}\left(s \mid N_c = n\right) &= \int_s^{\alpha} \frac{P\left(\Gamma_k = s, \Gamma_{n-1} = u, \Gamma_n > \alpha\right)}{P\left(\Gamma_{n-1} \le \alpha, \Gamma_n > \alpha\right)} \mathrm{d}u \\
&= \frac{1}{P\left(\Gamma_{n-1} \le \alpha, \Gamma_n > \alpha\right)} \int_s^{\alpha} P\left(\sum_{i=1}^{k} \alpha_i = s\right) P\left(\sum_{i=k+1}^{n-1} \alpha_i = u\right) P\left(\alpha_n > \alpha - u\right) \mathrm{d}u \\
&= \frac{1}{P\left(\Gamma_{n-1} \le \alpha, \Gamma_n > \alpha\right)} \int_s^{\alpha} (1 + u - \alpha) \frac{s^{k-1}}{(k-1)!} \frac{(u-s)^{n-k-2}}{(n-k-2)!} \mathrm{d}u \\
&= \frac{s^{k-1}}{(k-1)!} \left[\frac{\left(\left((\alpha - s)^{n-k-1} / (n-k-1)!\right) - \left((\alpha - s)^{n-k} / (n-k)!\right)\right)}{\left(\left(\alpha^{n-1} / (n-1)!\right) - \left(\alpha^n / n!\right)\right)} \right]
\end{aligned}
$$

We use this conditional density to compute $E(\Gamma_k \mid N_c = n)$ as follows:

$$
\begin{aligned}
E\left(\Gamma_k \mid N_c = n\right) &= \int_0^{\alpha} \frac{s^{k-1}}{(k-1)!} \left[\frac{\left(\left((\alpha - s)^{n-k-1} / (n-k-1)!\right) - \left((\alpha - s)^{n-k} / (n-k)!\right)\right)}{\left(\left(\alpha^{n-1} / (n-1)!\right) - \left(\alpha^n / n!\right)\right)} \right] \mathrm{d}s \\
&= k \frac{n!}{k!(n-k)!} \frac{1}{n\alpha^{n-1} - \alpha^n} \left[(n-k-a) \int_0^{\alpha} s^k (\alpha - s)^{n-k-1} \mathrm{d}s + \int_0^{\alpha} s^{k+1} (\alpha - s)^{n-k-1} \mathrm{d}s \right]
\end{aligned}
$$

Making a variable change $t = s/\alpha$, we see both integrals are beta functions and rearrange the expression

$$
\begin{aligned}
E\left(\Gamma_k \mid N_c = n\right) &= k \frac{n!}{k!(n-k)!} \frac{1}{n\alpha^{n-1} - \alpha^n} \left[(n-k-\alpha) \alpha^n B(k+1, n-k) + \alpha^{n+1} B(k+2, n-k) \right] \\
&= \frac{k}{(n-k)((n/\alpha) - 1)} \left[n - k - \alpha + \alpha \frac{k+1}{n+1} \right] = h_k(n, \alpha)
\end{aligned}
\tag{A4}
$$

Note that (A4) holds for $k < n - 1$ and $n \ge 3$. For $k = n - 1$ and $n \ge 2$

$$
f_{\Gamma_{n-1}}\left(s \mid N_c = n\right) = \frac{f_{\Gamma_{n-1}}(s) P\left(\alpha_n > \alpha - s\right)}{P\left(\Gamma_{n-1} \le \alpha, \Gamma_n > \alpha\right)}
$$

Using the above expression, we eventually obtain

$$
\begin{aligned}
E\left(\Gamma_{n-1} \mid N_c = n\right) &= \frac{1}{\left(\alpha^{n-1} / (n-1)!\right) - \left(\alpha^n / n!\right)} \int_0^{\alpha} \frac{s^{n-1}}{(n-2)!} (1 - \alpha + s) \mathrm{d}s \\
&= \frac{n(n-1)}{n\alpha^{n-1} - \alpha^n} \left[(1 - \alpha) \frac{\alpha^n}{n} + \frac{\alpha^{n+1}}{n+1} \right] = h(n, \alpha)
\end{aligned}
\tag{A5}
$$

APPENDIX D (Continued)

Equation (5) gives

$$
\begin{aligned}
E\left(C_{Z_{N_c}}\right) &= E\left(C^* E\left(\sum_{i=1}^{N_c} X_i - \sum_{i=2}^{N_c}\sum_{j=1}^{i-1} X_i\alpha_j \mid N_c = n\right)\right) \\
&= C^*\left[\sum_{n=2}^{\infty} P\left(N_c = n\right)\left\{\frac{n}{\lambda} - E\left(\sum_{i=2}^{n}\sum_{j=1}^{i-1} X_i\alpha_j \mid N_c = n\right)\right\} + \frac{P\left(N_c = 1\right)}{\lambda}\right] \\
&= C^*\left[\sum_{n=2}^{\infty} P\left(N_c = n\right)\left\{\frac{n}{\lambda} - \sum_{i=2}^{n} E\left(X_i\right) E\left(\Gamma_{i-1} \mid N_c = n\right)\right\} + \frac{(1-\alpha)}{\lambda}\right] \\
&= \frac{C^*}{\lambda}\left[\sum_{n=2}^{\infty} P\left(N_c = n\right)\left\{n - \sum_{i=2}^{n} E\left(\Gamma_{i-1} \mid N_c = n\right)\right\} + (1-\alpha)\right].
\end{aligned}
$$

Finally, substituting Equations (A4) and (A5), we obtain

$$
\begin{aligned}
E\left(C_{Z_{N_c}}\right) = \frac{C^*}{\lambda}\Bigg[&\sum_{n=3}^{\infty} P\left(N_c = n\right)\left\{n - \sum_{k=1}^{n-2} h_k(n,\alpha) - h(n,\alpha)\right\} \\
&+ P\left(N_c = 2\right)(2 - h(2,\alpha)) + (1-\alpha)\Bigg] = \frac{\bar{C}(\alpha)C^*}{\lambda}
\end{aligned}
$$

where $\bar{C}(\alpha)$ is defined as

$$
\bar{C}(\alpha) = \sum_{n=3}^{\infty} P\left(N_c = n\right)\left\{n - \sum_{k=1}^{n-2} h_k(n,\alpha) - h(n,\alpha)\right\} + P\left(N_c = 2\right)(2 - h(2,\alpha)) + (1-\alpha)\,.
$$

APPENDIX E:
PUBLICATION 4: A PREDICTIVE DECISION AID METHODOLOGY FOR DYNAMIC MITIGATION OF INFLUENZA PANDEMICS

In this appendix, we present the version of the manuscript "A Predictive Decision Aid Methodology for Dynamic Mitigation of Influenza Pandemics" currently in the second round of review in the special issue on "Optimization in Disaster Relief " in the OR Spectrum by Springer Publisher. The co-authors, Dr. Alex Savachkin, Alfredo Santana, Diana Prieto, and Dr. Tapas Das, authorized to include this document in my dissertation. Springer Publisher retains the copyright of this manuscript. The written authorization from the publisher to include the paper in my Ph.D. dissertation is attached in Appendix A.

APPENDIX E (Continued)

A Predictive Decision Aid Methodology for Dynamic Mitigation of Influenza Pandemics

Andrés Uribe-Sánchez
Alex Savachkin*
Alfredo Santana
Diana Prieto-Santa
Tapas K. Das
Department of Industrial and Management Systems Engineering
University of South Florida, Tampa, FL 33620

October 4, 2010

Abstract

The Institute of Medicine has recently pointed out the need for dynamic mitigation strategies for pandemic influenza. We present a large-scale simulation-based optimization methodology for developing dynamic predictive mitigation strategies for a network of regional pandemic outbreaks. The methodology incorporates varying virus epidemiology and population dynamics, and generates a progressive allocation of a limited resource budget to procure vaccines and antivirals, their administration capacities, and resources to enforce social distancing. The methodology considers measures of morbidity, mortality, and social distancing, translated into the societal and economic costs of lost productivity and medical expenses. The model was calibrated using historic pandemic data and implemented on a sample cross-regional outbreak in Florida, U.S., involving over four million inhabitants. We present a sensitivity analysis for estimating the marginal impact of changes in the total budget availability and variability of some critical mitigation parameters. The methodology is intended to assist public health policy makers.

Keywords: pandemic influenza, mitigation, dynamic, vaccination, antiviral, social distancing.

*Corresponding author: 4202 E. Fowler Ave., ENB118, Tampa, FL, USA, tel: +1-813-974-5577, e-mail: alexs@usf.edu

1 Motivation

Pandemic influenza (PI) outbreaks have historically entailed enormous societal calamities aggravated by tremendous economic forfeitures. Since 2003, scattered outbreaks of the avian-to-human transmittable H5N1 virus have been appearing through Asia, the Pacific region, Africa, Europe, and the Near East [1]. As of December 2009, the World Health Organization (WHO) has reported 447 confirmed H5N1 infected cases which resulted in 263 mortalities worldwide [2]. Although instances of H5N1 *human-to-human* transmission have rarely shown up in the recent statistics, ever-mutating emerging virus strains with unpredictable epidemiology pose a menacing threat to the humankind. A milder human-to-human transmissible H1N1 virus subtype resurfaced in Mexico in the Spring of 2009 and swiftly propagated to a global H1N1 pandemic outbreak. As of late December 2009, 208 countries have been affected with a total number of mortalities of at least 11,516 [3]. Most experts have an ominous expectation that the next pandemic will likely be caused by an emerging highly pathogenic virus subtype, to which there is little or no preexisting immunity in humans [4].

Combating influenza pandemic outbreaks requires an understanding of real-time evolution of virus epidemiology and population dynamics. The practicality and effectiveness of mitigation strategies also heavily depend on available emergency response infrastructure, mitigation resources, and allocation policies. At present, although the most commonly studied virus strain H5N1 has reached a worldwide case fatality ratio of more than 60% in recent years [2], prediction of the exact emerging virus subtype remains a challenging task. With modern technology, even after a new virus subtype is identified, a surge production of adequate stockpiles of a potent vaccine can take up to six months [5, 6]. This may prove to be forbiddingly long for the vaccine to be an effective mitigation resource in the critical onset stage of a pandemic. In the best case scenario, even if the emerged subtype has a known epidemiology, the existing stockpiles would be insufficient [7, 8]. Furthermore, supply of antiviral drugs, immunizers and other healthcare providers, hospital beds, resources to enforce social distancing, and logistics will also be substantially constrained. Hence, development of mitigation strategies must be done amidst limited knowledge of disease and population dynamics, constrained infrastructure, and limited availability and effectiveness of clinical treatments. This challenge, attested by the recent H1N1 outbreak, has been acknowledged by WHO [8] and echoed by national public health authorities, including the U.S. Centers for Disease Control and Prevention (CDC) [9, 10].

2 Status of Current Literature

The existing general literature on pandemic modeling aims to address various aspects of the pandemic process including (i) underlying spatio-temporal structure, (ii) contact mechanism, (iii) disease transmission, (iv) disease natural history, and (v) development of containment and mitigation strategies. These aspects are closely interrelated. For instance, the nature of the spatio-temporal structure, including composition of the social mixing groups and temporal dynamics of the affected population, drives the contact process which is the main determinant of the disease transmission [11–14]. A comprehensive decision-aid model for containment and mitigation must take into account all of the above constituents; it must incorporate the mechanism of disease progression, from initial infection, to asymp-

tomatic earlier phase, symptom manifestation, and final health outcome (recovery or death) [15–17]; it must also consider the population dynamics of disease spread, including individual susceptibility [18, 19], transmissibility [15, 20–22], and human behaviors that mediate infection generation [23–26]; finally, it must incorporate the impact of pharmaceutical and non-pharmaceutical prevention and intervention, including vaccination, antiviral therapy, social distancing, school and workplace closures, travel restrictions, and use of low-cost measures, such as face masks and hand washing [14, 17, 27–33]. In essence, effective mitigation strategies have a two-fold objective: (i) systemically to alter the disease dynamics and control disease progression with available clinical therapies, and (ii) to alter the social dynamics and contain disease propagation within the affected communities. Mitigation strategies vary in the composition of the target groups, geo-spatial coverage, and implementation time.

The current literature on assessment and development of PI containment and mitigation strategies can be broadly classified into (i) statistical models, (ii) mathematical models, (iii) simulation-based approaches, and (iv) combinations of thereof. In what follows, we present a summary survey of these approaches, mostly focusing on the simulation-based approaches.

The statistical models, driven mainly by likelihood or regression-based approaches, have primarily been used for epidemiological parameter assessment and estimation of the pandemic impact [22, 34, 35]. Traditionally, these models have inherently featured relatively simple and general spatio-temporal structures (e.g., homogeneous social mixing groups [21, 36–38]).

The mathematical models have mostly focused on modeling virus spread and policy assessment. Notable examples of such models are dynamic compartmental approaches, typically represented in the form of a set of differential equations, which delineate transitions between disease phases (e.g., susceptible, exposed, symptomatic infected, etc.) [11, 39–41]. Based on the solution approach, the mathematical models can be subdivided into analytical (or closed form) and iterative. Compared to the statistical models, the dynamic-iterative models feature more granular composition of the mixing groups [11, 33, 42]. However, the degree of granularity is still limited since any additional spatio-temporal considerations can negatively impact the computational robustness of the models. Description of the contact processes in mathematical models generally does not take into account changes in the behavioral patterns during the course of the pandemic (e.g., compliance to intervention by vaccination and social distancing) [39–41]. Consideration of behavioral aspects can be found in some of the recent simulation-based models [43–46]. Furthermore, mathematical models are typically based on the infection pathways and disease progression that are invariant to time and individual attributes [40, 41].

The simulation-based approaches have been used for modeling virus spread and assessment and generation of pharmaceutical and/or non-pharmaceutical interventions. Based on the way of generating disease progression, these models can be categorized as those that track infection pathway of each individual entity [13, 27, 47–51] and the rest that are driven by occurrence of infection events [14, 31]. In contrast to the statistical and mathematical models, the simulation models are capable of providing most detailed description of population dynamics whereby each individual can be assigned a set of attributes (e.g., age, gender, community, etc), that can be modified without altering the general model structure [47, 52]. However, such comprehensive descriptive granularity is achieved at the expense of higher data demand and substantial computational burden. As a result, most existing simulation-based

approaches incorporate statistical and/or mathematical submodels (e.g., for infection generation) in order to attain an effective balance between model accuracy and practicality [47, 51].

In recent years, the simulation-based models have focused on integration of therapeutical and non-therapeutical prevention and intervention, to develop synergistic strategies aimed at result-oriented use of constrained resources. These approaches first aim to implement a form of social distancing to reduce the contact between the susceptible and the infected. The infected population is then treated with an antiviral therapy to reduce infectiousness, and the susceptible are vaccinated to increase their immunity. For example, [14] implemented social distancing for all contacted and symptomatic cases followed by antiviral application. Such strategies, which appear to be more discriminating and thus less expensive, have been found to be particularly efficient for low transmissibility scenarios (i.e., scenarios with the values of the basic reproduction number below 1.8 [48, 53]).

Most notable among recent efforts is a 2006-2007 initiative by MIDAS, the Models of Disease Infectious Agent Study network, which studied three independent simulation models [54]. These models were used to emulate large-scale PI spread for rural areas of Asia [48, 55], U.S. and U.K. [13, 49], and the city of Chicago [56]. MIDAS cross-validated the models by simulating the city of Chicago, with 8.6M inhabitants, and implementing targeted layered containment (TLC). Under TLC, the symptomatic infectious cases refrain from going to work (take liberal leave), receive antiviral treatment, and become subject to household quarantine; the asymptomatic contacts receive targeted antiviral prophylaxis (TAP) and become subject to household quarantine. The research findings of the MIDAS network and other institutions [17, 31] were used in a recent "Modeling Community Containment for Pandemic Influenza" report by the Institute of Medicine (IOM), U.S., to formulate a set of recommendations for mitigating PI at the local level [57]. These recommendations were used in a pandemic preparedness guidance developed jointly by CDC, HHS, and other federal U.S. agencies [58].

The IOM report [57] points out several limitations of the MIDAS models, observing that "*because of the significant constraints placed on the models*" being considered by policy makers, "*the scope of models should be expanded.*" The IOM recommends that "*steps be taken now to adapt or develop decision-aid models that can be readily linked to surveillance data to provide real-time feedback during an epidemic. Research protocols should be developed, approved, and put in place now to generate the information needed during an outbreak to inform models, and improve their disease sub-models.*" The report also strongly recommends 1) "*that future modeling efforts incorporate broader outcome measures ... to include the costs and benefits of intervention strategies*", and 2) "*that models examining the potential effectiveness of school and workplace closures on mitigating pandemic influenza include a broader range of closure options.*" We add here that most current approaches focus on assessment of policies defined *apriori*; few of the existing models for design of synergistic mitigation strategies are *"learning"*, i.e., capable of predicting of and adapting to changes in the course of a pandemic, and ultimately generating a dynamic optimal strategy [51]. Furthermore, the majority of the current simulation models feature a single-region design (see Appendix, Table 1). In [51], we developed a simulation-based optimization model for generating mitigation strategies for cross-regional pandemic outbreaks. The model was implemented on a sample outbreak and the resultant strategy was compared to the existing governmental pro-rata distribution policy, which allocates mitigation resources to each

affected region in proportion to its population [59].

In this paper, we present a novel decision-aid methodology for developing dynamic predictive mitigation strategies for a network of regional pandemic outbreaks. The methodology is driven by a large-scale simulation-based dynamic optimization model that incorporates varying virus epidemiology and region-specific population dynamics. The model generates mitigation strategies for an efficient, progressive allocation of limited resources, including stockpiles of vaccines and antiviral drugs, healthcare capacities for administration of vaccination and antiviral therapy, and social distancing enforcement resources. The optimization control seeks dynamically to minimize the impact of ongoing outbreaks and the expected impact of potential outbreaks, and allocates the resources accordingly. The methodology considers measures of morbidity, mortality, and social distancing, translated into the cost of lost productivity and medical expenses (societal and economic costs). The model was calibrated using historic pandemic data and tested on a sample cross-regional outbreak in Florida, U.S., with over four million inhabitants. We also present a sensitivity analysis for estimating the marginal impact (measured in terms of the average total pandemic cost and the average number of fatalities) of changes in the total budget availability and variability of some critical decision factors. These factors included: (i) vaccine efficacy, (ii) efficacy of antiviral therapy, (iii) social distancing conformance level, (iv) social distancing declaration threshold, and (v) social distancing period.

Compared to our previous work in [51], this paper features the following main advances: (i) in [51], progressive resource allocations are irrevocable, i.e., once resources are allocated to an affected region, they remain in the region until full depletion, regardless of the posterior dynamics of the overall pandemic; in contrast, our model is capable of re-allocating the resources remaining from the previous distributions, based on the current pandemic status, and thus achieving a more efficient resource utilization; (ii) our model incorporates the cost of the resources (e.g., vaccines, antiviral, etc.) and strives to allocate a total available budget, as opposed to a separate allocation of total available quantities of individual resources, which vary significantly in their relative cost and effectiveness; (iii) our model investigates optimal policy generation under two scenarios of virus severity: low transmissibility and high transmissibility, as opposed to a high transmissibility analysis in [51]; (iv) our study attempts to analyze the affect of social distancing policies (namely, the target population conformance) on the dynamics of societal and economic costs; (v) in this paper, we also present a short description of our decision-aid simulation software made freely available to general public through our website.

This paper has the following organization. In Section 3, we present our simulation-based optimization methodology, including description of the population dynamics and disease transmission, the mechanism of disease progression, and therapeutical and non-therapeutical intervention, followed by description of the calibration methodology for single-region and cross-regional simulation models, and presentation of the optimization control model. In Section 4, results of the testbed implementation are presented, followed by discussion of the sensitivity analysis. Conclusions are given in Section 5.

3 Simulation-Based Optimization Model

Our large-scale simulation-based optimization model generates predictive strategies to allocate a total available budget of mitigation resources over a network of regional pandemic outbreaks, progressively, from one affected region to the next. Mitigation resources include stockpiles of vaccine(s) and antiviral drug(s), hospital beds, capacities for vaccination and antiviral administration, social distancing enforcement resources, among others. The methodology combines a cross-regional simulation model, a set of single-region simulation models, and an overarching dynamic optimization control.

The regions inside the network are classified as unaffected, ongoing outbreak (which includes new outbreak), and contained (see Fig. 1). The regions are interconnected by air and land travel, which is emulated by the cross-regional model. The single-region model mimics the population and disease dynamics inside each ongoing region, impacted by available pharmaceutical and non-pharmaceutical prevention and intervention. The pandemic spreads from ongoing to unaffected regions by infectious travelers who pass through regional border control. At every new regional outbreak episode (epoch), the cross-regional model invokes the optimization control, which allocates the total available resource budget, including remaining resources from the previous allocations, to the new/ongoing outbreak regions (*actual allocation*) and potential (unaffected) outbreak regions (*virtual allocation*). The objective function of the optimization model incorporates measures of morbidity, mortality, and social distancing, translated into the cost of lost productivity and medical expenses. The objective function strives to minimize the total cost of the new/ongoing outbreaks and the expected cost of the potential outbreaks, spreading from the ongoing regions. Detailed daily pandemic statistics are collected for each affected region, including the numbers of new infected, deceased, and quarantined cases, for different age groups. As the regional outbreaks become contained, the model estimates their actual societal and economic costs.

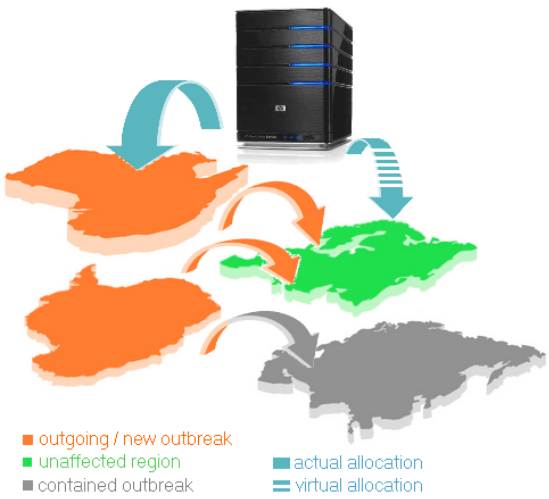


Figure 1: Schematic of simulation-based optimization methodology

In the remainder of this section, we present the details of the cross-regional simulation model (Section 3.1) and the single-region simulation model (Section 3.2), including the description of the calibration methodology (Section 3.3). The dynamic optimization control model is presented in Section 3.4, followed by analysis of the results of the testbed implementation in Section 4.

3.1 Cross-Regional Simulation Model

The cross-regional simulation model emulates propagation of the pandemic across the network of affected regions. It controls a set of single-region simulation models of ongoing outbreaks and invokes the optimization model for actual and virtual resource allocation at every new outbreak epoch. The

schematic of the cross-regional simulation model is presented in Fig. 2.

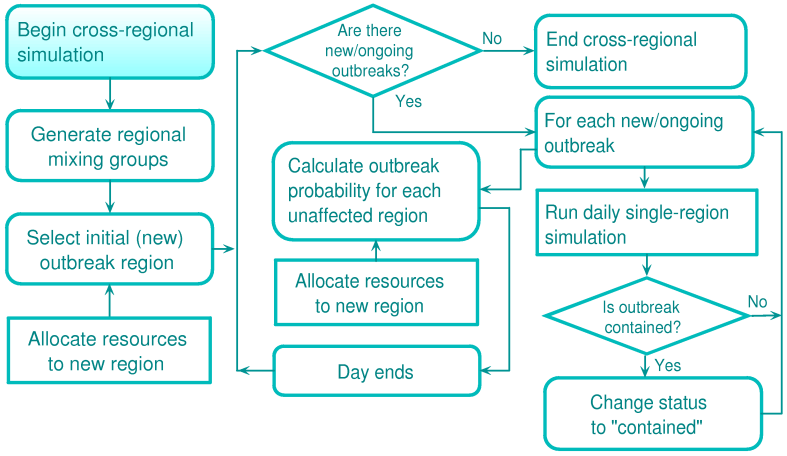


Figure 2: Schematic of cross-regional simulation model

The model is initialized by generating mixing groups and population dynamics for each network region (for details, see Section 3.2.1). A pandemic is triggered by injecting one or more infectious cases into a randomly selected region, designated as the initial outbreak region. Details associated with the resulting contact dynamics and disease propagation within the region are presented in Section 3.2. As the symptomatic cases start seeking medical assistance, the new regional outbreak is detected. At this point, the model calls the optimization control which generates a resource allocation (see Section 3.4). The cross-regional simulation then passes control back to the single-region model, which executes a cycle (e.g., daily) of the regional disease and population dynamics, now mediated by the allocated clinical therapies and social distancing measures (see Section 3.2). The simulation clock inside the single-region routine advances in smaller time increments (e.g., hourly; see Section 4).

As the outbreak intensifies, it spreads over to the unaffected regions, as infectious travelers pass undetected through air and land border control with some probability (the probabilities are different for asymptomatic and symptomatic cases). Travelers are considered to act independently. Each infectious traveler is assumed to initiate a regional outbreak with an equal, time-homogeneous probability ω for the entirety of his/her infection period, regardless of his/her point of origin. For each unaffected region, the outbreak probability at time $t > 0$, P_t , is calculated using the binomial probability law, as follows

$$P_t = 1 - (1 - \omega)^{n_t} \ , \tag{1}$$

where n_t denotes the number of infectious travelers in the region at time t. Based on the outbreak probability value, the cross-regional model determines which of the unaffected regions have become new outbreaks (in the testbed implementation, the values of P_t were computed once at the end of each day). The model also determines if an outbreak has been contained, based on a certain threshold of the daily infection rate. The cross-regional simulation ends when all outbreaks have been contained.

3.2 Single-Region Simulation Model

The single-region simulation model emulates population and disease dynamics within an affected region.

A schematic of the model is shown in Fig. 3. The model subsumes the following main components: (i) population dynamics, (ii) contact and infection process, (iii) disease natural history, and (iv) mitigation prevention and intervention, including measures of social distancing, vaccination, and antiviral application. The model collects detailed regional influenza statistics, including numbers of infected, recovered, deceased, and quarantined cases, for different age groups. For a contained outbreak, its societal and economic costs are calculated. The societal cost includes the cost of lost lifetime productivity of the deceased; the economic cost includes the cost of medical expenses of the recovered and deceased and the cost of lost productivity of the quarantined (see Section 4.1.3). The single-region simulation model builds upon a prototype presented in [47].

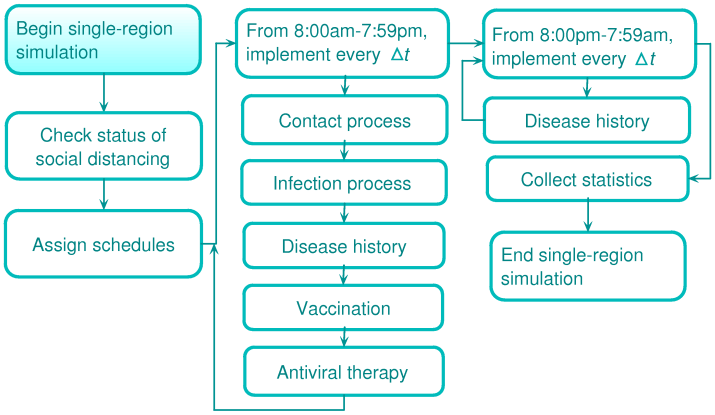


Figure 3: Schematic of single-region simulation model

At any point of time, the population of an ongoing region is assumed to be composed of the following exclusive compartments: susceptible, contacted, infected, recovered, and deceased (see Fig. 4). During the course of his/her social interaction, a susceptible individual may periodically come into contact with infectious cases. The contacted individual then either becomes infected with a certain probability p or returns to the compartment of susceptibles. An infected case then either dies with a certain probability m or recovers. It is further assumed that a recovered person develops immunity and cannot be susceptible again. All recovered cases continue circulating through the mixing groups. In what follows, we present the details of the main components of the single-region simulation model.

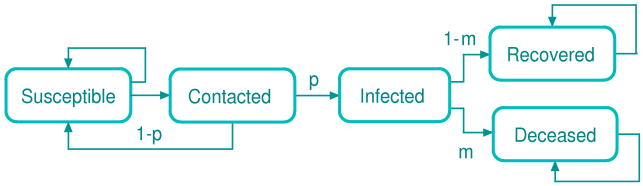


Figure 4: Schematic of the disease progression model

3.2.1 Population Dynamics

We model a region as a set of population centers composed of *mixing groups* of various types, including households, offices, manufacturing facilities, universities, schools, churches, shopping centers, entertainment centers, etc. A household consists of household members, each of which is assigned a comprehensive set of attributes including age, gender, parenthood, workplace, immunity status, infection susceptibility, probability of travel, and others. Each inhabitant is also assigned Δt time-discrete

(e.g., $\Delta t = 1$ hour) weekday and weekend schedules, which depend on a number of factors, including: (i) age and parenthood, (ii) inhabitant's disease status, (iii) travel status, (iv) social distancing decrees in place, and (v) inhabitant's conformance to them. As their schedules advance, inhabitants circulate throughout the mixing groups, staying a certain amount of time in each of them.

3.2.2 Contact and Infection Processes

Disease transmission occurs through contact events between infectious and susceptible individuals within the mixing groups. At the beginning of each period Δt (e.g., one hour), for each mixing group g, the simulation model tracks the total number of infectious cases n_g present in the group. Each infectious case randomly generates r_g per Δt unit of time *new contacts*, uniformly, from the susceptible cases present in the mixing group. We assume the following simplifying characterization of the contact process: (i) during Δt period, a susceptible may come into contact with at most one infectious case and (ii) each contact exposure lasts Δt units of time. Once a susceptible has started a contact exposure at time t, he/she will develop into an infectious case at time $t + \Delta t$ with a certain probability that is calculated as shown below.

Let $L_i(t)$ be a nonnegative continuous random variable that represents the duration of contact exposure, starting at time t, that is required for a contact i to become infected. We assume that $L_i(t)$ follows an exponential distribution with parameter $\lambda_i(t)$, where $\lambda_i(t)$ represents the instantaneous force of infection applied to contact i at time t [42, 60, 61]. The probability that a susceptible i, whose contact exposure has started at time t, will develop into an infectious case at time $t + \Delta t$ is then given as

$$P\{L_i(t) \leq \Delta t\} = 1 - e^{-\lambda_i(t)\Delta t}. \tag{2}$$

3.2.3 Disease Natural History

It is assumed that upon becoming infected, an individual enters simultaneously into the phases of latency and incubation (see Figure 5). During the incubation phase, the individual stays asymptomatic (i.e., shows no visible symptoms). At the end of the latency phase, the individual becomes infectious and enters the infectious phase [49, 54, 55, 62]. At the end of the incubation period, the individual becomes symptomatic. At the conclusion of the infectious period, the individual enters the final disease stage which culminates in his/her recovery or death.

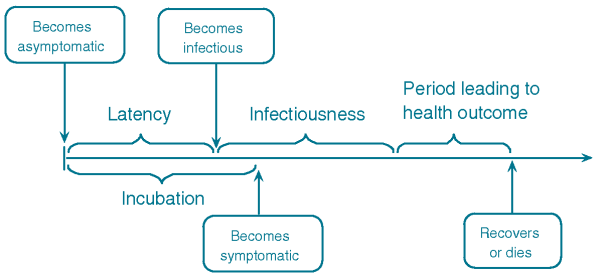


Figure 5: Schematic of disease natural history model

Mortality for influenza like diseases is a complex process, which is affected by a number of factors and variables. For most of these variables, little or no accurate data has been collected from past pandemics. In addition, the time of death could oftentimes be weeks following the disease episode (mainly attributable to subsequent pneumonia related complications [63]). Because of the uncertainty underlying the mortality process, we adopted a simplified, age-based form of the mortality probabilities,

where the mortality probability of infected i, m_i is given as

$$m_i = \mu_i - \tau\rho_i, \tag{3}$$

where μ_i is the age-dependent base mortality probability of infected i, ρ_i is his/her status of antiviral treatment (0 or 1), and τ is the efficacy of the antiviral therapy, measured in terms of the reduction in the base mortality probability [55].

3.2.4 Mitigation Strategies

Mitigation prevention and intervention considered in the single-region model include pharmaceutical and non-pharmaceutical measures. Implementation of the mitigation measures is initiated upon detection of the first confirmed infected case [64]. At this point, mitigation resources are assigned (see Section 3.4) and deployed in the region. The model considers a certain outbreak detection delay and a delay for deployment of field responders.

Pharmaceutical mitigation. Pharmaceutical prevention and intervention consist of vaccination and antiviral application. Vaccines are offered to the individuals from a prespecified risk group, to reduce their infection susceptibility. It is assumed that a certain fraction of the risk group will not comply with vaccination. Once administered, the vaccine takes a certain period to become effective (typically, between 10 and 14 days) [65]. Vaccination is constrained by the available stockpile and the administration capacity, measured in terms of the number of immunizer-hours.

For antiviral application, we assume that a certain fraction of symptomatic infected cases will seek medical assistance [66, 67]. Those of them who belong to a prespecified mortality risk group, receive an antiviral treatment, to reduce their mortality probability (see Eq. 3). It is assumed that an antiviral becomes effective immediately. The antiviral application is subject to availability of the antiviral stockpile and the administration capacity, measured in terms of the number of certified nurses. Both the antiviral application and vaccination are affected by a number of social behavioral factors, including conformance of the target population, its degree of risk perception, and associated compliance of healthcare personnel [68–70]. The conformance level of the target population could be affected by the demographical profile of the region [71–75] and quality of the public information available [76], among other factors. The degree of risk perception of the target population could be impacted by a negative experience developed during pharmaceutical campaigns of the previous outbreaks [77, 78], as well as by public fear and rumors [79, 80].

Non-pharmaceutical mitigation. For a social distancing mediation, we adopted a guidance suggested by CDC, U.S. [58]. The guidance establishes five categories of pandemic severity (from 1 to 5) and recommends different quarantine and closure policies for each of the categories. The categories are determined based on the value of the case fatality ratio (CFR), the cumulative proportion of the number of fatalities in the total infected population. Our simulation model periodically reassesses the CFR value during the pandemic course. For the CFR values lower than 0.1% (which corresponds to Category 1), voluntary at-home isolation of the infected cases is *implemented*. If the CFR falls in the range between 0.1% and 1.0% (Categories 2 and 3), in addition to the above at-home isolation, the following measures are *recommended*: (i) voluntary quarantine of households members of infected

cases and (ii) child and adult social distancing. Finally, for the CFR values exceeding 1.0% (Categories 4 and 5), all the above measures are *implemented*. Alike pharmaceutical measures, the effectiveness of social distancing mitigation is also affected by several of the behavioral factors mentioned above [76]. Our model considers a certain social distancing conformance level, which can vary based on the demographics profile. Travel restrictions considered in the model included regional air and land border control for outgoing infected travelers (for details, see the testbed implementation in Section 4).

3.3 Calibration Methodology

The single-region simulation model was calibrated using two commonly accepted measures of pandemic severity: the *basic reproduction number* (R_0) and the *infection attack rate* (IAR). R_0 is defined as the average number of secondary infections, produced by a typical infected case in a totally susceptible population. Our model with its a detailed, person-to-person infection generation traceability allows identification of all secondary infections created by each infected case. All infected cases are then classified by generation of infection, as in [31, 48], where a generation is defined as the set of all infected cases (offsprings) that are at the same tier of descent from their infection generators (ancestors) [61]. The value of R_0 is then calculated as the average reproduction number of a typical generation in the early stage of the pandemic when no interventions are implemented (known as the *baseline* scenario). Our model was calibrated to attain the baseline values of R_0 similar to those obtained from historic pandemic data [48, 53] (see Section 4.1.2).

IAR is defined as the ratio of the total number of infections over the pandemic period to the size of the initial susceptible population. To further calibrate our model, we used the following relationship between baseline R_0 and IAR, as presented in [14, 61]:

$$R_0 = \frac{-\ln(1 - IAR)}{IAR}, \quad \textit{for } R_0 \geq 1,\ 0 < IAR < 1. \tag{4}$$

Section 4.1.2 of this paper provides the details of the calibration process for a sample testbed scenario.

3.4 Optimization Model

The optimization model is invoked at every new outbreak epoch to allocate the total available resource budget, including remaining resources from the previous allocations, to the new/ongoing outbreak regions (*actual allocation*) and potential outbreak regions (*virtual allocation*). By doing so, the model seeks to progressively minimize the impact of ongoing outbreaks and the expected impact of potential outbreaks. Mitigation resources include stockpiles of vaccine(s) and antiviral drug(s), hospital beds, capacities for vaccination and antiviral administration, social distancing enforcement resources, among others. The objective function of the optimization model incorporates measures of morbidity, mortality, and social distancing, translated into the cost of lost productivity and medical expenses. In what follows, we present the details of the model. We introduce the following notation.

S = the set of all regions,

A^n = the set of regions in which pandemic is contained at the n^{th} outbreak epoch ($n = 1, 2, \dots$),

B^n = the set of ongoing regions at the n^{th} outbreak epoch,

C^n = the set of unaffected regions at the n^{th} outbreak epoch,
M^n = budget availability at the n^{th} outbreak epoch,
R = the set of available types of mitigation resources ($R = \{1, 2, \ldots, r\}$),
c_i = unit cost of type i resource, $i \in R$,
$\hat{q}_i^{\,n}$ = amount of resource i remaining from previous allocations, at the n^{th} outbreak epoch.

Let TC_k^n denotes the total cost of an outbreak in region k at the n^{th} outbreak epoch. The total cost is a function of the *decision variables* q_{ik}^n which denote the amount of resource i allocated to region k at the n^{th} outbreak epoch. We assume that outbreaks occur one at a time. At the n^{th} new outbreak epoch in region j, the following optimization problem is invoked

$$Min \quad TC_j^n(q_{1j}^n, q_{2j}^n, \ldots, q_{rj}^n) + \sum_{l \in B^n \setminus \{j\}} TC_l^n(q_{1l}^n, q_{2l}^n, \ldots, q_{rl}^n) + \sum_{s \in C^n} TC_s^n(q_{1s}^n, q_{2s}^n, \ldots, q_{rs}^n) \cdot p_s^n$$

subject to

$$\sum_{i \in R} c_i \cdot q_{ij}^n + \sum_{l \in B^n \setminus \{j\}} \sum_{i \in R} c_i \cdot q_{il}^n + \sum_{s \in C^n} \sum_{i \in R} c_i \cdot q_{is}^n \cdot p_s^n - \sum_{i \in R} c_i \cdot \hat{q}_i^{\,n} \leq M^n$$

$$q_{ij}^n + \sum_{l \in B^n \setminus \{j\}} q_{il}^n + \sum_{s \in C^n} q_{is}^n \cdot p_s^n \geq \hat{q}_i^{\,n}, \quad \forall i \in R.$$

In the above objective function, the first term represents the total cost of the new outbreak in region j, estimated at the n^{th} outbreak epoch, and based on the actual resource allocation $\{q_{1j}^n, q_{2j}^n, \ldots, q_{rj}^n\}$. The second term represents the total cost of ongoing outbreaks, excluding region j, which is (re)estimated at the n^{th} outbreak epoch, based on the current pandemic status (for details, see below). This cost is a function of the allocation $\{q_{1l}^n, q_{2l}^n, \ldots, q_{rl}^n\}$, which may include amounts remaining from previous allocations. The third term represents the total expected cost of outbreaks in currently unaffected regions, based on the virtual allocation $(q_{1s}^n, q_{2s}^n, \ldots, q_{rs}^n)$ and the regional outbreak probabilities p_s^n.

The first model constraint relates the total available budget with the value of the current actual and virtual resource allocations, adjusted with the value of the resources remaining from the previous allocations. The second set of constraints guarantee that the needs of the current actual and virtual allocations will first be satisfied using the resources remaining from the previous allocations (the outstanding resource needs will then be fulfilled from the remaining budget).

The total cost of an outbreak in region k at the n^{th} outbreak epoch is calculated as following

$$TC_k^n = \sum_{h \in \mathcal{H}} (m_h + \bar{w}_h) X_{hk}^n + \sum_{h \in \mathcal{H}} \bar{w}_h \cdot Y_{hk}^n + \sum_{h \in \mathcal{H}} \hat{w}_h \cdot D_{hk}^n + \sum_{h \in \mathcal{H}} w_h \cdot V_{hk}^n + \sum_{h \in \mathcal{H}} \hat{m}_h \cdot U_{hk}^n \,, \qquad (5)$$

where

$\mathcal{H}$ = the set of age groups,
m_h = total medical cost of an infected case in age group h over his/her disease period,
$\hat{m}_h$ = total medical cost of an uninfected case in age group h over the pandemic period,
$\bar{w}_h$ = total cost of lost wages of an infected case in age group h over his/her disease period,
$\hat{w}_h$ = cost of lost lifetime wages of a deceased case in age group h,
w_h = daily cost of lost wages of a non-infected case in age group h who complies with quarantine,
X_{hk}^n = total number of infected cases in age group h who seek medical assistance,

Y^n_{hk} = total number of infected cases in age group h who do not seek medical assistance,
D^n_{hk} = total number of deceased cases in age group h,
V^n_{hk} = total number of person-days of cases in age group h who comply with quarantine,
U^n_{hk} = total number of uninfected cases in age group h.

Variables X^n_{hk}, Y^n_{hk}, D^n_{hk}, V^n_{hk}, and U^n_{hk} are defined for region k at the n^{th} outbreak epoch. We determine the value of X^n_{hk} using the following regression model

$$X^n_{hk} = \delta^0_{hk} + \sum_{i \in R} \delta^i_{hk} \cdot q_{ik} + \sum_{i,m \in R,\, i \neq m} \delta^{im}_{hk} \cdot q_{ik} \cdot q_{mk} \ , \tag{6}$$

where $\delta^i_{\cdot\cdot}$ denotes the regression coefficient associated with resource i, and $\delta^{im}_{\cdot\cdot}$ is the regression coefficient for the interaction between resources i and m. Similar expressions are used for Y_{hk}, D_{hk}, and V_{hk}. U^n_{hk} is obtained by subtracting X^n_{hk} from the total population of the region at the n^{th} outbreak epoch.

We have that $p^n_k = \sum_{l \in B^n} p^n_{lk}$, where p^n_{lk} denotes the outbreak probability in region k, caused by an ongoing outbreak in region l, estimated at the n^{th} outbreak epoch. This probability is considered to be a function of the resource allocation for region l at the n^{th} outbreak epoch, and is calculated using the following regression model

$$p^n_{lk} = \gamma^0_{lk} + \sum_{i \in R} \gamma^i_{lk} \cdot q_{il} + \sum_{\substack{i,m \in R \\ i \neq m}} \gamma^{im}_{lk} \cdot q_{il} \cdot q_{ml} \ , \tag{7}$$

where $\gamma^i_{\cdot\cdot}$ denotes the regression coefficient associated with resource i, $\gamma^{im}_{\cdot\cdot}$ is the regression coefficient associated with interaction between resources i and m, and $\gamma^0_{\cdot\cdot}$ represents the intercept.

3.5 Simulation Optimization Algorithm

Below we present the algorithm for the simulation optimization model.

1. Estimate regression equations for all regions using the single-region simulation model (Eq. 6, 7).
2. Set $n = 1$. Initialize sets of regions: $A^n = \emptyset$, $B^n = \emptyset$, $C^n = S$.
3. Select randomly the initial outbreak region j.
4. Update sets of regions: $B^n \leftarrow B^n \cup \{j\}$ and $C^n \leftarrow C^n \backslash \{j\}$.
5. Solve the resource allocation model for region j. Update the total budget availability.
6. If $B^n \neq \emptyset$, do Step 7. Else, do Step 9.
7. (a) For each ongoing region, implement a next day run of its single-region simulation.
   (b) Check the containment status of each ongoing region. Update sets A^n and B^n, if needed.
   (c) For each unaffected region, calculate its outbreak probability.
   (d) Based on the outbreak probability values, determine if there is a new outbreak region(s) j. If there is no new outbreak(s), go to Step 6. Otherwise, go to Step 8.
8. For each new outbreak region j,

(a) Increment $n \leftarrow n + 1$.

(b) Update sets $B^n \leftarrow B^n \cup \{j\}$ and $C^n \leftarrow C^n \backslash \{j\}$.

(c) Re-estimate regression equations for each region $k \in B^n \cup C^n$ using the single-region simulations, where each simulation is initialized to the current outbreak status in the region.

(d) Determine the remaining availability of the previously allocated resources.

(e) Solve the resource allocation model for each region $k \in B^n$.

(f) Update the total budget availability.

9. Calculate the total cost (Eq. 5) for each contained region and update the overall pandemic cost.

4 Testbed Illustration

A sample cross-regional outbreak scenario included a network of four counties in Florida, U.S.: Hillsborough, Miami Dade, Duval, and Leon, with populations of 1.0, 2.2, 0.8, and 0.25 million people, respectively. The H5N1 virus subtype was considered. A basic unit of time Δt for people's schedules, contact dynamics, infection transmission, disease natural history, and implementation of interventions was taken to be one hour (see Fig. 3 and Eq. 2). Each regional simulation was run for a period (up to 185 days) until the number of new daily infections approached near zero (see Section 4.1.3).

4.1 Model Parameter Values

This section presents the details on selecting parameter values for cross-regional and single-region models, including parameters of population and disease dynamics, calibration, and mitigation.

4.1.1 Parameters of Population and Disease Dynamics

Demographic and social dynamics data for each of the regions was extracted from the U.S. Census [81] and the National Household Travel Survey [82, 83] (see Appendix, Tables 2-4). Compositions of the regional mixing groups are shown in Appendix, Tables 5-8 [83]. In these tables, column 1 and 2 show the mixing group type and the number of mixing groups for each type, respectively. Column 3 shows the probability distribution among the types for assigning the workplaces of inhabitants. Columns 4-6 show the probability distribution among the types for assigning the weekday after-work errands, weekend errands, and errands during quarantine (see § 4.1.3), respectively. The last column contains the hourly contact rates for each mixing group type. The hourly schedules [83] were adopted from [47].

Each infected individual was assigned a daily travel probability of 0.24% [82], of which 7% was by air and 93% by land transportation. The origin-destination travel probabilities within the network of four regions were calculated based on the traffic volume data [84–87] (see Appendix, Table 9). Infection detection probabilities for regional border controls for symptomatic cases were assumed to be 95% and 90% [88], for air and land transportation, respectively. These values were reduced by 70% for asymptomatic travelers [88]. Travel bans were implemented for all detected infectious cases. An undetected infectious traveler was assumed to trigger a regional outbreak in his/her destination with a

time-homogeneous probability $\omega = 0.001$ during his/her infectiousness period. The regional outbreak probabilities P_t were calculated once at the end of each day using Eq. 1.

The instantaneous force of infection applied to contact i at time t ($\lambda_i(t)$ in Eq. 2) was modeled as

$$\lambda_i(t) = -ln(1 - p_i(t)), \text{ where } p_i(t) = \alpha_i - \delta\theta_i(t), \tag{8}$$

where α_i is the age-dependent base instantaneous infection probability of contact i, θ_i(t) is his/her status of vaccination at time t (taken as 0 or 1), and δ is the vaccine efficacy, measured in terms of the reduction in the base instantaneous infection probability (achieved after 10 days [65]). Note that, as $p_i(t)$ increases, the instantaneous force of infection $\lambda_i(t)$ grows at an increasing rate (see Fig. 6). The age-dependent base instantaneous infection probabilities were adopted from [49] (see Appendix, Table 10).

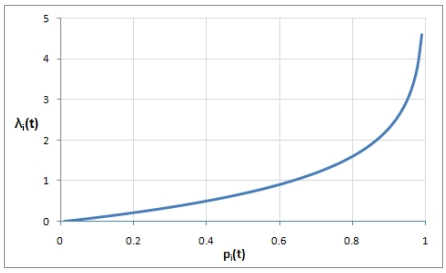


Figure 6: $\lambda_i(t)$ vs. $p_i(t)$

The disease natural history for H5N1 virus subtype was taken as the following: a latent period of 29 hours, an incubation period of 46 hours, and an infectiousness period between 29 and 127 hours [89].

The values of the base mortality probability of infected case i, μ_i (Eq. 3), were determined using the statistics recommended by the Working Group on Pandemic Preparedness and Influenza Response [90]. This data shows the percentage of mortality for age-based high-risk groups (see Appendix, Table 11, columns 1-3). For the testbed scenario, the mortality probabilities for different age groups (Appendix, Table 11, column 4) were estimated using the assumption that high-risk cases are expected to account for 85% of the total number of fatalities, for each age group [90].

4.1.2 Calibration of the Single-Region Models

The single-region simulation models were calibrated against R_0 and IAR (see Section 3.3), using hourly contact rates within mixing groups. Original contact rates were adopted from [49]. These contact rates were adjusted to obtain the baseline values of R_0 similar to those estimated from past outbreaks [48, 53], for both high and low transmissibility scenarios (see Appendix, Table 12 for a sample of contact rates).

The values of R_0 were estimated for each region using the average reproduction numbers for generations of infection, as presented in [31, 48], over multiple replicates (e.g., see Fig. 7 for the Hillsborough County). As the figure illustrates, generations 5 through 8 (the solid line) and generations 7 to 9 (the dotted line) represent "typical" [31] or representative infectious cases, born in the *early stage* of the pandemic, when most of the regional population is susceptible. For the purpose of computing the value of R_0, earlier generations were disregarded, as they were composed of a limited number of infected cases with highly variable individual reproduction numbers.

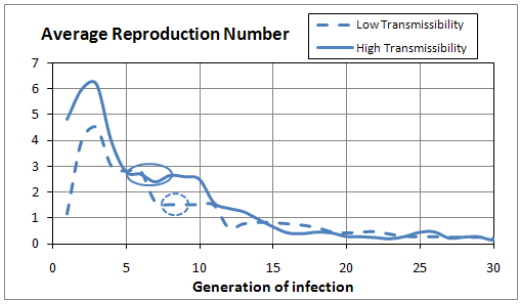


Figure 7: Estimation of R_0 (Hillsborough)

Historically, R_0 values for PI range between 1.4 and 3.9, where PI with $R_0 \leq 1.8$ are considered as of low transmissibility and PI with $2.2 \leq R_0 \leq 3.9$ as of high transmissibility [48, 53]. Henceforth, the R_0 value of 2.538 for the high transmissibility testbed (Fig. 7, solid line) and 1.525 for the low transmissibility testbed (Fig. 7, dotted line) were taken, for the Hillsborough County (other regions had similar R_0 values). These values corresponded to the simulated $IARs$ of 0.881 and 0.538, respectively. Note that from the theoretical relationship in Eq. 4, IAR of 0.881 and 0.538 yield R_0 of 2.426 and 1.435, respectively. Thus, the simulated R_0 numbers show a good match with the theoretical approximations.

4.1.3 Mitigation Related Parameters

The mitigation resources considered in the testbed included stockpiles of vaccines and antiviral, administration capacities for vaccination and antiviral therapy, and quarantine enforcement resources (required to achieve a targeted social distancing conformance level). We assumed a 24-hour CDC resource deployment delay once the first infection case is confirmed [64].

Pharmaceutical measures. The vaccination risk group included individuals younger than 5 years and older than 65 years [91]. The risk group for antiviral application included individuals below 15 years and above 55 years [91, 92]. The efficacy levels for the vaccine, δ (in Eq. 8) and antiviral, τ (in Eq. 3) were assumed to be 40% [55, 93] and 30%, respectively. See Section 4.2.2 for a sensitivity analysis on both parameters. We assumed a 95% target population conformance for both measures. The immunity development period for the vaccine was taken as 10 days [65]; the antiviral was assumed to become effective immediately. Table 13 summarizes vaccination and antiviral treatment resource requirements for each region along with resource costs [94–96] and the total budget requirement.

Non-pharmaceutical measures. A simplified version of the CDC guidance for non-pharmaceutical interventions for Category 5 was implemented (see Section 3.2.4, [58]). Once the case fatality ratio has reached 1.0% in the affected region, a social distancing policy is declared and remains in effect for 14 days [58]. Individuals below a prespecified age ξ (22 years) were assumed to stay at home during the quarantine period. Of the remaining population, a certain proportion ϕ [97] stayed at home and was allowed a one-hour leave, every three days, to buy essential supplies. The remaining $(1 - \phi)$ non-compliant proportion followed a regular schedule. All testbed scenarios considered the quarantine conformance level ϕ (a decision variable) bounded between 50% and 80% [76, 97].

An outbreak was considered contained, if the daily new infections did not exceed five cases, for seven consecutive days. Once contained, a region was simulated for an additional 10 days [98] to allow an accurate estimation of the pandemic statistics. The costs of lost productivity and medical expenses were adopted from [90] and inflation-adjusted using [99] (see Appendix, Table 14). The medical costs of uninfected arising from the use of face masks and preventive medicine were not considered.

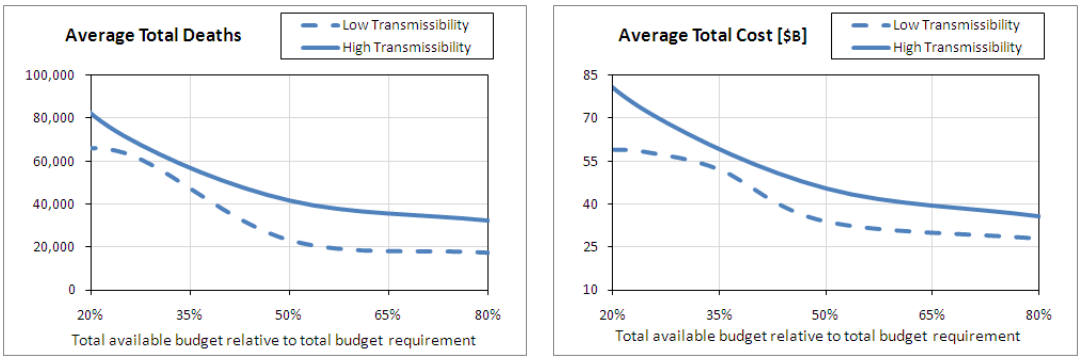
The optimization model was based on regression equations re-estimated at every outbreak epoch. For each region, we developed a 2^5 statistical design of experiment, to estimate the regression coefficients of the significant decision variables (factors) and their interactions. To ease the very significant computational burden, the testbed implementation considered allocation decisions only for new outbreak regions. The simulation code was written in C++ and run on a Pentium IV with a 3.40 GHz CPU and 4.0 GB RAM. The running time for a cross-regional simulation replicate averaged 20 minutes.

4.2 Sensitivity Analysis

This section presents a sensitivity analysis for assessing the marginal impact of changes in the total budget availability and variability of some of the critical factors, for both low and high transmissibility scenarios. The marginal impact was measured separately by the change in the total pandemic cost and the number of mortalities (averaged over multiple replicates), resulting from a unit change in the total budget availability or a factor value, one at a time. Factors under consideration included: (i) antiviral efficacy, (ii) vaccine efficacy, (iii) social distancing conformance, (iv) social distancing declaration threshold, and (v) social distancing period. We also investigated the affect of varying the social distancing conformance on the dynamics of the societal and economic costs.

4.2.1 Total Budget Availability

Figures 8(a) and 8(b) show the dynamics of the pandemic impact, measured as the average number of fatalities and total cost, for different levels of the total budget availability relative to the total budget requirement (see Appendix, Table 13).



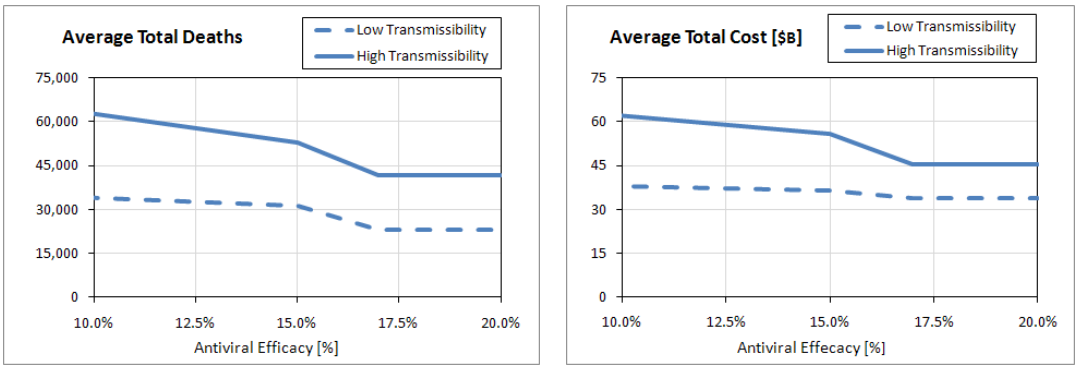
(a) (b)

Figure 8: Sensitivity analysis on total budget availability (measured in terms of the average number of deaths (a) and the average total cost (b))

As expected, the curves for the average number of deaths and total cost exhibit a downward trend, for both transmissibility scenarios, as the total budget availability increases. An increased budget translates into higher availabilities of mitigation resources, which will mediate regional pandemic impact and reduce the probability of spread to unaffected regions. As also expected, a higher virus transmissibility generates more infections which, in turn, result into more fatalities and, subsequently, bigger societal and economic costs. As the budget availability approaches the budget requirement (starting from approximately 60%), both impact curves show a converging behavior, for both scenarios, whereby the marginal impact of additional resource availability decreases. This can be explained by noting that the total budget requirement was calculated assuming the worst case scenario where *all* regions are affected and provided with adequate resources to cover their respective populations at risk.

4.2.2 Antiviral Efficacy

Figures 9(a) and 9(b) show the behavior of the two impact measures for the level of the antiviral efficacy (τ in Eq. 3) between 10% and 20%, and a fixed level of the total budget availability (50% of the total requirement).



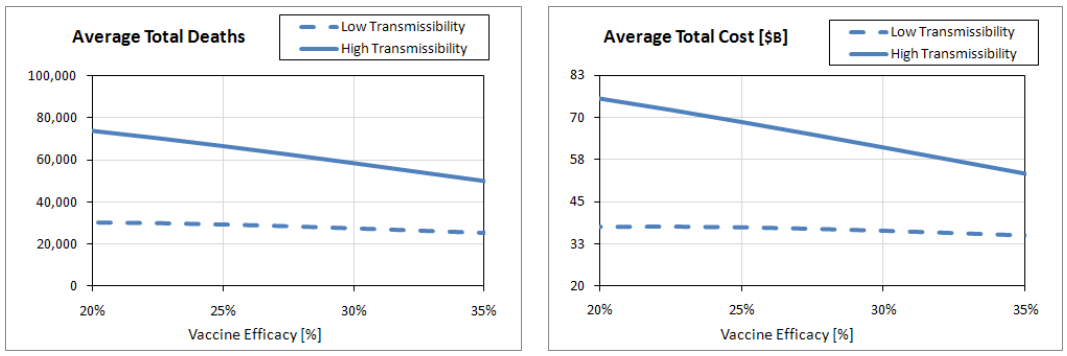
(a) (b)

Figure 9: Sensitivity analysis on antiviral efficacy (measured in terms of the average number of deaths (a) and the average total cost (b))

As expected from Eq. 3, for both transmissibility scenarios, the two curves exhibit a decreasing trend which is approximately linear in the range of τ between 10% and 15%. As the value of τ approaches the value of the maximum base mortality probability (approximately 16% for the elderly risk group), the resultant effective mortality probability tends to zero, which explains the converging behavior of the curve representing the total number of mortalities. The total cost curve exhibits a similar pattern.

4.2.3 Vaccine Efficacy

Both impact measure curves exhibit a downward trend as the vaccine efficacy (δ in Eq. 8) increases between 20% and 35%, with a fixed level of the total budget availability (50% of the total requirement)



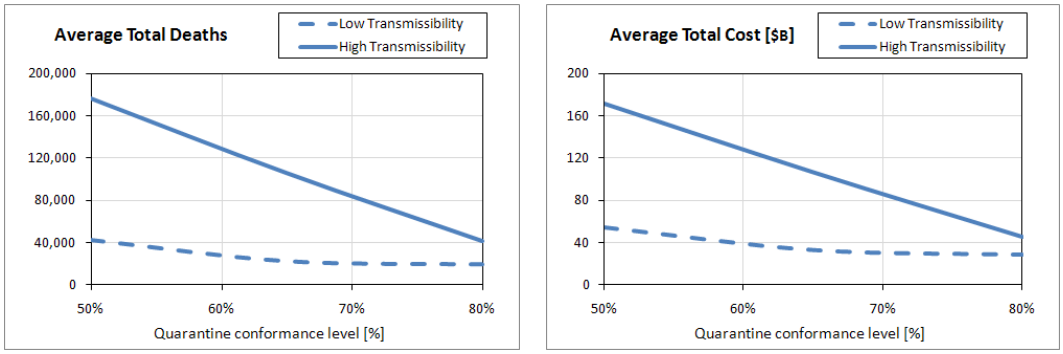
(a) (b)

Figure 10: Sensitivity analysis on vaccine efficacy (measured in terms of the average number of deaths (a) and the average total cost (b))

(see Fig. 10(a) and 10(b)). For lower transmissibility, the marginal utility of the vaccine efficacy is less significant which can be explained by noting that such scenarios generate fewer infections and consequently, the overall impact of a more effective vaccine is less pronounced.

4.2.4 Social Distancing Conformance Level

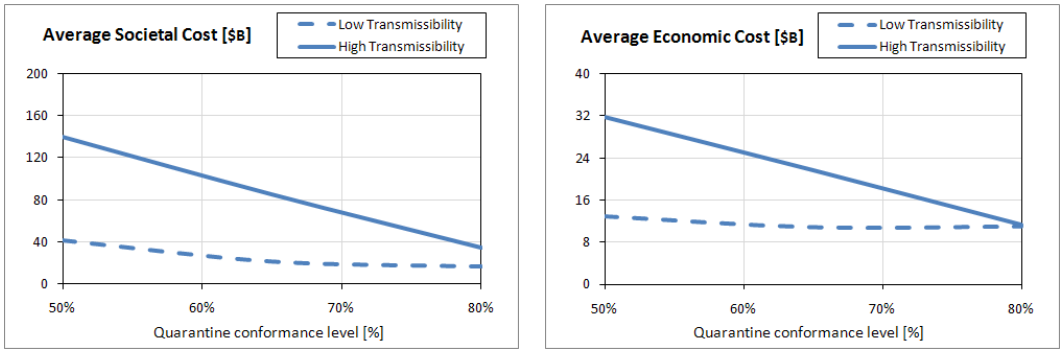
Reduction of the contact intensity through social distancing has long proven to be one of the most efficient mitigation mechanisms. Figures 11(a) and 11(b) show the dynamics of the average number of fatalities and total cost for different levels of the social distancing conformance. The analysis was conducted for conformance levels between 50% and 80%.



(a) (b)

Figure 11: Sensitivity analysis on social distancing conformance (measured in terms of the average number of deaths (a) and the average total cost (b))

For both transmissibility scenarios, the two curves exhibit a downward trend, which can be attributed to a reduced contact intensity associated with higher conformance. The trends are steeper for higher transmissibility scenarios which are characterized by more intensive social dynamics within the mixing groups. The reduction in the contact intensity gets amplified throughout generations of infection within the affected region and, more importantly, leads to reduced probabilities of spread to unaffected regions.



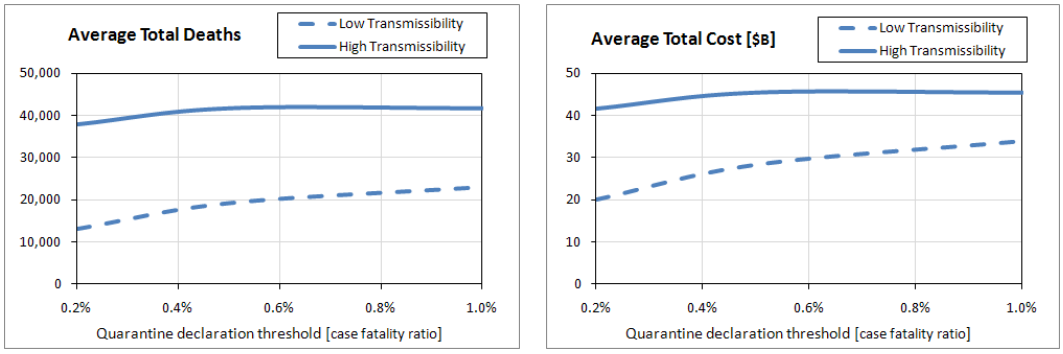
(a) (b)

Figure 12: Sensitivity analysis on social distancing conformance (societal (a) and economic cost (b))

Figures 12(a) and 12(b) show the dynamics of the total cost broken into the societal and economic components. The societal cost for the two scenarios generally decreases with quarantine conformance as a consequence of generating fewer infections and deaths (and hence, smaller lost productivity) during the quarantine period. For higher transmissibility, the marginal impact of the conformance level is more pronounced due to the effect of amplifying reduction in contact intensity explained above. A similar behavior can be observed for the economic cost which incorporates the cost of medical expenses of the recovered/deceased (over the entire pandemic period) and the cost of lost productivity of the quarantined individuals (during the social distancing period).

4.2.5 Social Distancing Declaration Threshold

Figures 13(a) and 13(b) show the dynamics of the impact measures for different levels of the social distancing declaration threshold, measured in terms of the CFR (see Section 4.1.3). The analysis was conducted for the CFR values between 0.2% and 1.0%.



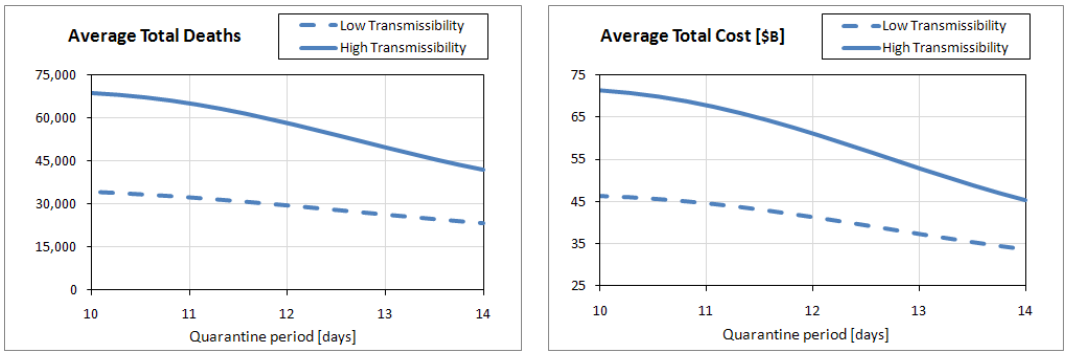
(a) (b)

Figure 13: Sensitivity analysis on social distancing declaration threshold (measured in terms of the average number of deaths (a) and the average total cost (b))

For both scenarios, the trends are increasing since later declaration leads to a growth in the number of infected and dead. It can also be observed that for higher transmissibility, the curves reach saturation starting from CFR of approximately 0.55%. This can be explained by noting that in this case, the time difference in quarantine declaration using CFR values between 0.55% and 1.0% is insignificant.

4.2.6 Social Distancing Period

Figures 14(a) and 14(b) show the dynamics of the average number of fatalities and total cost for values of the social distancing period between 10 and 14 days [58]. Similar to the analysis for the social distancing conformance level (see Section 4.2.4), for both transmissibility scenarios, the two curves exhibit a decreasing trend: once the case fatality ratio reaches a significant value of 1.0%, social distancing becomes the most efficient containment measure. From this point, any additional quarantine days (starting from 10 and up to 14 total days [58]) will reduce both the contact intensity during the quarantine period and also the size of the post-quarantine infectious population.



(a) (b)

Figure 14: Sensitivity analysis on social distancing period (measured in terms of the average number of deaths (a) and the average total cost (b))

Based on our methodology, we have also developed a decision-aid simulator with a GUI which is made freely available to general public through our web site at `http://imse.eng.usf.edu/pandemics.aspx`. The simulator allows the input of data for regional demographic and social dynamics, and disease related parameters (see Appendix, Fig. 15). It is intended to assist public health decision makers in conducting customized what-if analysis for assessment of mitigation options and development of policy guidelines. Examples of such guidelines include targeted risk groups for vaccination and antiviral treatment, social distancing policies (e.g., thresholds for declaration and lifting, closure options (i.e., household-based, schools, etc.), and compliance targets), and guidelines for travel restrictions.

5 Conclusions

The decision-aid methodology presented in this paper incorporates varying virus epidemiology and region-specific population dynamics. The model supports development of mitigation strategies for an efficient, progressive allocation of a limited resource budget over a network of regional outbreaks. The model seeks to dynamically minimize the impact of ongoing outbreaks and the expected impact of potential outbreaks, spreading from the ongoing regions. The methodology considers measures of morbidity, mortality, and social distancing, translated into the societal and economic costs of lost productivity and medical expenses.

As our analysis shows, for both transmissibility scenarios, starting with the budget availability of approximately 60%, the marginal impact of additional resources steadily decreases. This can be explained by observing that while the total budget requirement was determined assuming that *all* regions would be affected, our dynamic predictive allocation generally decreases the probability of spread from the ongoing regions, which in turn reduces the actual resource need. The analysis also shows that compared to high transmissibility, the marginal utility of the vaccine efficacy for low transmissibility is less significant due to a relatively smaller infected population, which makes the overall impact of a more effective vaccine less pronounced. We have observed that the overall pandemic cost is significantly affected by the social distancing conformance, particularly for higher transmissibility scenarios charac-

terized by more intense social interactions. A higher degree of social distancing conformance leads to a reduction in the contact intensity within the affected region. This further reduces the probabilities of spread to unaffected regions. We have also observed that later declaration of social distancing leads to a growth in the number of infected and dead. Moreover, for both transmissibility scenarios, longer social distancing period (from 10 to 14 days [58]) significantly reduces the pandemic cost by decreasing both the contact intensity and the size of the post-quarantine infectious population.

To the best of our knowledge, the presented methodology is one of the first attempts to offer a *dynamic predictive* decision-aid tool, which incorporates measures of both *societal and economic impact*. Compared to the existing models, such as [51], this model makes the following contributions: (i) the model is capable of re-allocating the resources remaining from the previous distributions, based on the current pandemic status, and thus achieves a more efficient resource utilization; (ii) the model incorporates the costs of the resources and aims to allocate a total available budget, as opposed to allocating available quantities of individual resources, which vary in their relative cost and effectiveness; (iii) the testbed implementation considers optimal policy generation under both low and high transmissibility scenarios; (iv) finally, the paper attempts to investigate the dynamics of the societal and economic costs under varying social distancing policies. In addition, the simulation represents one of the first of its kind in incorporating a broader range of social behavioral aspects. The simulation model features a flexible design which can be particularized to an even broader range of pharmaceutical and non-pharmaceutical interventions and more granular mixing groups. The developed methodology is intended to assist public health decision makers in development of dynamic strategies for mitigation of large-scale cross-regional pandemic outbreaks.

6 Appendix

Author, year	Single region (SR) / Cross regional (CR)	Model objective	Key features
Ferguson et al., 2005 [48], Ferguson et al., 2006 [13]	SR (Thailand, 2005), SR (U.S. & UK, 2006)	Model PI spread & assess mitigation strategies	- 85M Thailand, 300M US, 58.1M UK - Use of GIS (Landscan) - Use of targeted mass prophylaxis and social distancing - A set of homogeneous mixing groups
Glass et al., 2006 [31]	SR (small town in New Mexico, U.S.)	Examine role of social distancing	- Targeted social distancing - Fixed, small-scale contact network - Time between infection events follows an exponential distribution
Germann et al., 2006 [49]	SR (U.S.)	Assess mitigation strategies	- 281M inhabitants (2000 U.S. census data) divided in 2000-person communities - Sensitivity analysis on R_0 from 1.6 to 2.4
Wu et al., 2006 [14]	SR (Hong Kong)	Test different intervention scenarios	- Natural history includes pre-symptomatic cases - Use of household-based interventions - Need for significant stocks of antiviral drugs
Colizza et al., 2007 [23, 100]	CR (global)	Model worldwide spread of a pandemic	- Use of an air travel network - Diverse urban centers - Use of compartmental models (SLIR) - Analysis of antiviral and travel restrictions
Halloran et al., 2008 [54]	SR (Chicago)	Cross-validate targeted layered containment models (Ferguson/Germann/Eubank)	- 8.6M people - R_0 from 1.9 to 3.0 - Assessment of intervention strategies
Cooley et al., 2007 [101]	SR (North Carolina, U.S.)	Compare a simulated pandemic curve against real-life data	- Use of 2003-04 North Carolina data - Use of ILI data to estimate model parameters
Das, Savachkin & Zhu, 2008 [47]	SR	Mimic stochastic propagation of PI and assess mitigation strategies	- Large-scale model (1.1M) - Detailed schedules for inhabitants - Heterogeneous mixing groups
Savachkin et al., 2010 [51]	CR (Florida, U.S.)	Model PI spread & assess comprehensive dynamic mitigation strategies	- Dynamic predictive large-scale simulation-based optimization methodology - 4M people testbed - Use of vaccination, prophylaxes, and social distancing
STEM-Eclipse, 2009 [102]	CR (global)	Model worldwide PI spread	- Geographic visualization of PI spread - Use of SIR model - Limited to land transportation

Table 1: A summary of simulation-based PI containment and mitigation models

	Adult Population Distribution by Age			
Age Group \ Region	**Hillsborough**	**Miami Dade**	**Duval**	**Leon**
23 - 29	0.16	0.15	0.16	0.24
30 - 64	0.67	0.66	0.69	0.63
65 - 99	0.17	0.19	0.15	0.13

Table 2: Distribution of regional adult population by age [81]

	Children Population Distribution by Age			
Age Group \ Region	**Hillsborough**	**Miami Dade**	**Duval**	**Leon**
0 - 5 (pre-school)	0.24	0.22	0.24	0.16
6 - 9 (elementary school)	0.23	0.23	0.25	0.17
10 - 14 (middle school)	0.25	0.25	0.23	0.17
15 - 17 (high school)	0.13	0.14	0.14	0.10
18 - 22 (college)	0.15	0.16	0.14	0.40

Table 3: Distribution of regional children population by age [81]

Household Type		Regional Population by Household Type			
# Adults	**# Children**	**Hillsborough**	**Miami Dade**	**Duval**	**Leon**
1	0	0.28	0.25	0.27	0.30
1	1	0.04	0.04	0.04	0.04
2	0	0.31	0.26	0.30	0.32
1	2	0.04	0.05	0.05	0.04
2	1	0.13	0.15	0.14	0.13
1	3	0.01	0.01	0.01	0.01
2	2	0.13	0.15	0.13	0.11
1	4	0.01	0.01	0.01	0.00
2	3	0.06	0.08	0.06	0.04

Table 4: Distribution of regional population by households [81]

Mixing Group (MG) Type	Number of MG	Distribution of Workplaces	Distribution of Weekday Errands	Distribution of Weekend Errands	Distribution of Quarantine Errands	Hourly Contact Rate
Home	1	0.066	0.000	0.000	0.800	1.500
Factory	613	0.058	0.000	0.000	0.000	0.750
Office	2,266	0.302	0.000	0.000	0.000	0.750
Pre-school	224	0.005	0.000	0.000	0.000	1.050
Elementary school	66	0.010	0.000	0.000	0.000	2.573
Middle school	134	0.203	0.000	0.000	0.000	3.750
High school	59	0.097	0.000	0.000	0.000	3.750
College	46	0.106	0.000	0.000	0.000	3.750
Afterschool center	256	0.007	0.000	0.000	0.000	1.500
Grocery store	390	0.026	0.619	0.515	0.100	0.375
Restaurant	223	0.087	0.278	0.256	0.000	0.375
Entertainment center	360	0.032	0.066	0.116	0.000	0.375
Church	86	0.001	0.037	0.113	0.100	0.375

Table 5: Composition of mixing groups, Hillsborough County [82]

APPENDIX E (Continued)

Mixing Group (MG) Type	Number of MG	Distribution of Workplaces	Distribution of Weekday Errands	Distribution of Weekend Errands	Distribution of Quarantine Errands	Hourly Contact Rate
Home	1	0.092	0.000	0.000	0.800	1.500
Factory	1,353	0.035	0.000	0.000	0.000	0.750
Office	2,880	0.128	0.000	0.000	0.000	0.750
Pre-school	188	0.010	0.000	0.000	0.000	1.050
Elementary school	246	0.188	0.000	0.000	0.000	2.573
Middle school	84	0.098	0.000	0.000	0.000	3.750
High school	82	0.116	0.000	0.000	0.000	3.750
College	59	0.206	0.000	0.000	0.000	3.750
Afterschool center	507	0.006	0.000	0.000	0.000	1.500
Grocery store	942	0.025	0.619	0.515	0.100	0.375
Restaurant	3,935	0.085	0.278	0.255	0.000	0.375
Entertainment center	758	0.011	0.066	0.116	0.000	0.375
Church	266	0.000	0.037	0.113	0.100	0.375

Table 6: Composition of mixing groups, Miami Dade County [82]

Mixing Group (MG) Type	Number of MG	Distribution of Workplaces	Distribution of Weekday Errands	Distribution of Weekend Errands	Distribution of Quarantine Errands	Hourly Contact Rate
Home	1	0.049	0.000	0.000	0.800	1.500
Factory	519	0.063	0.000	0.000	0.000	0.750
Office	2,880	0.313	0.000	0.000	0.000	0.750
Pre-school	74	0.006	0.000	0.000	0.000	1.050
Elementary school	116	0.170	0.000	0.000	0.000	2.572
Middle school	35	0.083	0.000	0.000	0.000	3.750
High school	30	0.087	0.000	0.000	0.000	3.750
College	21	0.112	0.000	0.000	0.000	3.750
Afterschool center	245	0.006	0.000	0.000	0.000	2.000
Grocery store	320	0.024	0.619	0.515	0.100	1.500
Restaurant	1,474	0.074	0.278	0.255	0.000	0.375
Entertainment center	244	0.012	0.066	0.116	0.000	0.375
Church	77	0.001	0.037	0.113	0.100	0.375

Table 7: Composition of mixing groups, Duval County [82]

Mixing Group (MG) Type	Number of MG	Distribution of Workplaces	Distribution of Weekday Errands	Distribution of Weekend Errands	Distribution of Quarantine Errands	Hourly Contact Rate
Home	1	0.072	0.000	0.000	0.800	1.50
Factory	103	0.012	0.000	0.000	0.000	0.750
Office	1,093	0.212	0.000	0.000	0.000	0.750
Pre-school	20	0.008	0.000	0.000	0.000	1.050
Elementary school	30	0.106	0.000	0.000	0.000	2.573
Middle school	15	0.051	0.000	0.000	0.000	3.750
High school	14	0.064	0.000	0.000	0.000	3.750
College	9	0.374	0.000	0.000	0.000	3.750
Afterschool center	60	0.005	0.000	0.000	0.000	1.500
Grocery store	52	0.021	0.619	0.515	0.100	0.375
Restaurant	512	0.069	0.278	0.255	0.000	0.375
Entertainment center	73	0.006	0.066	0.116	0.000	0.375
Church	16	0.002	0.037	0.113	0.100	0.375

Table 8: Composition of mixing groups, Leon County [82]

	Inter-Regional Travel Probability			
Origin \ Destination	**Hillsborough**	**Miami Dade**	**Duval**	**Leon**
Hillsborough	0.00	0.60	0.27	0.13
Miami Dade	0.74	0.00	0.16	0.10
Duval	0.61	0.29	0.00	0.10
Leon	0.52	0.31	0.17	0.00

Table 9: Inter-regional travel probabilities [84–87]

Age Group	0-5	6-19	20-29	31-65	66-99
α_i	0.156	0.106	0.205	0.195	0.344

Table 10: Instantaneous infection probability for different age groups; adopted from [49]

Age Group	% High-Risk Cases	% Death for High-Risk Cases	Mortality Probability
0-19	6.4	9.0	0.007
20-64	14.4	40.9	0.069
65+	40.0	34.4	0.162

Table 11: Mortality probability for different age groups [90]

Mixing group	Hourly contact rate
Home	1.50
Factory	0.75
Office	0.75
Preschool	1.05
Elementary school	2.57
Middle school	3.75
High school	3.75
University	3.75
Afterschool center	1.50
Grocery store	0.38
Restaurant	0.38
Entertainment center	0.38
Church	0.38

Table 12: Hourly contact rates by mixing group (Hillsborough County, high transmissibility scenario); adopted from [49]

Region (population)	Resource Requirements by Region: Hillsb. (1,007,916)	Miami D. (2,209,702)	Duval (852,168)	Leon (248,761)	Total (4,318,547)	Cost of Resource	Required Budget by Resource
Resource							
Vaccine stock	305,036	679,181	241,522	76,007	1,301,745	\$8.48/dose	\$11,038,800
Antiviral stock	415,294	749,058	460,393	105,307	1,730,052	\$60/dose	\$103,803,140
No. nurses (antiv.)	650	1,104	786	166	2,706	\$27/hr	
						8 hr/day, 50 days	\$29,226,975
No. nurses (vacc.)	1,059	2,358	839	264	4,520	\$27/hr	
						8 hr/day, 14 days	\$13,668,326
					Total Budget Requirement		\$157,737,241

Table 13: Regional resource and budget requirements [94–96]

Pandemic Impact Measure (age group, years)	Value US$
Average cost of lost lifetime productivity of a deceased case (0 - 19)	$1,336,347.86
Average cost of lost lifetime productivity of a deceased case (20 - 64)	$1,370,987.28
Average cost of lost lifetime productivity of a deceased case (65 - 99)	$98,959.24
Average cost of lost productivity and medical expenses of a recovered/deceased case (0 -19)	$5,078.48
Average cost of lost productivity and medical expenses of a recovered/deceased case (20 -64)	$10,466.68
Average cost of lost productivity and medical expenses of a recovered/deceased case (65 -99)	$11,566.09
Average daily cost of lost productivity of a non-infected quarantined case (20-99)	$432.54

Table 14: Values of pandemic impact measures (societal and economic costs) [90, 99]

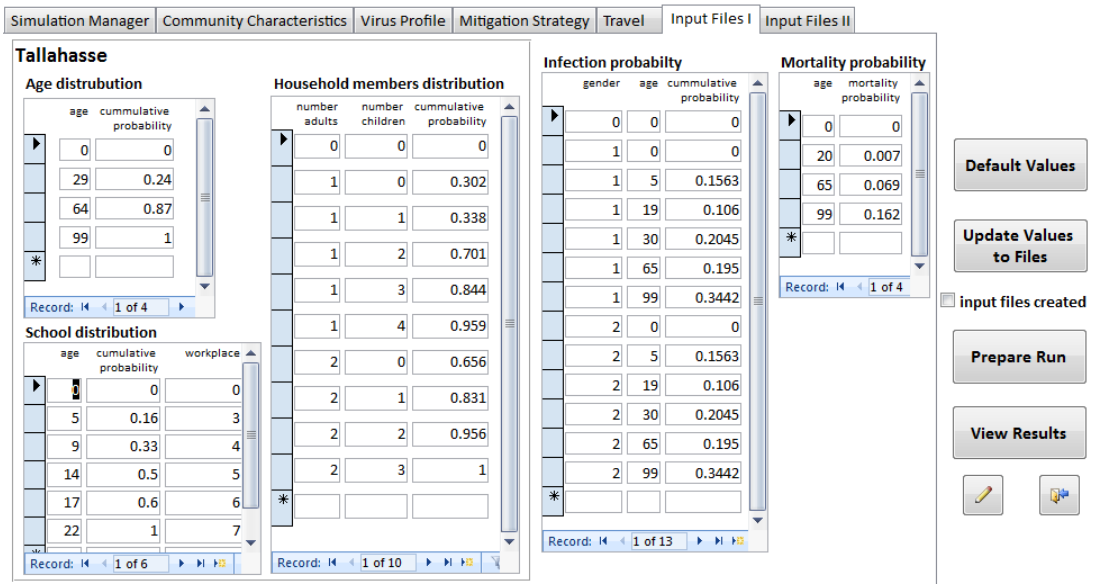


Figure 15: A snapshot of the decision-aid simulator GUI

Acknowledgments

We would like to acknowledge with thanks the many helpful suggestions made by Prof. Yiliang Zhu, Department of Epidemiology and Biostatistics at the University of South Florida, Tampa, FL, USA.

References

[1] Centers for Disease Control and Prevention (CDC). Avian influenza: Current H5N1 situation. `http://www.cdc.gov/flu/avian/outbreaks/current.htm#clusters`, 2008. Last accessed on 02/09/2008.

[2] World Health Organization (WHO). Cumulative number of confirmed human cases of Avian Influenza A(H5N1) reported to WHO. `http://www.who.int/csr/disease/avian_influenza/country/cases_table_2010_07_05/en/index.html`, 2010.

[3] World Health Organization. Pandemic (H1N1) 2009 - update 107. `http://www.who.int/csr/don/2010_07_02/en/index.html`, 2010. Last accessed on 07/08/2010.

[4] Arthur Schoenstadt. Spanish Flu. `http://flu.emedtv.com/spanish-flu/spanish-flu.html`, 2009. Last accessed on 12/15/2009.

[5] D.S. Fedson and S. Hant. Pandemic influenza and the global vaccine supply. *Clinical Infectious Deseases*, 36:1552–1561, 2003.

[6] J.G. Aunins, A.L. Lee, and D.B. Volkin. *Vaccine production*. Boca Raton: CRC Press LLC, 2 edition, 2000.

[7] World Health Organization. Pandemic (h1n1) 2009 vaccine deployment update - 17 december 2009. `http://www.who.int/csr/disease/swineflu/vaccines/h1n1_vaccination_deployment_update_20091217.pdf`, 2009.

[8] WHO Global Influenza Programme. Pandemic Influenza Preparedness and Response. Technical report, World Health Organization, 2009.

[9] Centers for Disease Control and Prevention (CDC). Preparing for Pandemic Influenza. `http://www.cdc.gov/flu/pandemic/preparedness.htm`, 2007. Last accessed on 04/27/2009.

[10] U.S. Department of Health & Human Services. HHS pandemic influenza plan. `http://www.hhs.gov/pandemicflu/plan/`, 2007. Last accessed on 03/27/2009.

[11] R. Larson. Simple models of influenza progression within a heterogeneous population. *Operations Research*, 55(399-412):165–195, 2007.

[12] M. E. Halloran, I. M. Longini, A. Nizam, and Y. Yang. Containing bioterrorist smallpox. *Science*, 298: 1428–1432, 2002.

[13] N. Ferguson, D. A. Cummings, C. Fraser, J. Cajka, Philip Cooley C., and D. S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(27):448–452, 2006.

[14] J. T. Wu, S. Riley, C. Fraser, and G. Leung. Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Med*, 3(9):1532–1540, 2006.

[15] A. Handel, I. Longini, and R. Antia. Towards a quantitative understanding of the within-host dynamics of influenza A infections. *Epidemics*, 1(3):185–195, 2009.

[16] B. Pourbohloul, A. Ahued, B. Davoudi, R. Meza, L.A. Meyers, D.M. Skowronski, I. Villaseñor, F. Galván, P. Cravioto, D.J.D. Earn, et al. Initial human transmission dynamics of the pandemic (H1N1) 2009 virus in North America. *Influenza Other Respi Viruses*, 3(5):215–222, 2009.

[17] M. Atkinson and L. Wein. Quantifying the Routes of Transmission for Pandemic Influenza. *Bulletin of Mathematical Biology*, 70(3):820–867, 2008.

[18] V.E. Pitzer, S.J. Olsen, C.T. Bergstrom, S.F. Dowell, and M. Lipsitch. Little evidence for genetic susceptibility to influenza A (H5N1) from family clustering data. *Emerg Infect Dis*, 13(7):1074–1076, 2007.

[19] V.E. Pitzer, G.M. Leung, and M. Lipsitch. Estimating variability in the transmission of severe acute respiratory syndrome to household contacts in Hong Kong, China. *American journal of epidemiology*, 166 (3):355–359, 2007.

[20] Y. Yang, J.D. Sugimoto, M.E. Halloran, N.E. Basta, D.L. Chao, L. Matrajt, G. Potter, E. Kenah, and I.M. Longini Jr. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science*, 326(5953): 729–733, 2009.

[21] Y. Yang, M.E. Halloran, J. Sugimoto, and I.M. Longini Jr. Detecting human-to-human transmission of avian influenza A (H5N1). *Emerg Infect Dis*, 13(9):1348–1353, 2007.

[22] S. Cauchemez, F. Carrat, C. Viboud, A. Valleron, and P. Boelle. A Bayesian (MCMC) approach to study transmission of influenza: application to household longitudinal data. *Statist. Med.*, 23:3469–3487, 2004.

[23] V. Colizza, A. Barrat, M. Barthelemy, and A. Vespignani. The role of the airline transportation network in the prediction and predictability of global epidemics. *PNAS*, 103:2015–2020, 2006.

[24] J.M. Epstein, J. Parker, D. Cummings, and R.A. Hammond. Coupled contagion dynamics of fear and disease: mathematical and computational explorations. *PLoS One*, 3(12), 2008.

[25] J.M. Epstein, D.M. Goedecke, F. Yu, R.J. Morris, D.K. Wagener, and G.V. Bobashev. Controlling pandemic flu: the value of international air travel restrictions. *PLoS One*, 2(5), 2007.

[26] M.E. Halloran. Invited commentary: Challenges of using contact data to understand acute respiratory disease transmission. *American Journal of Epidemiology*, 164(10):945, 2006.

[27] I. M. Longini, M. E. Halloran, A. Nizam, and Y. Yang. Containing pandemic influenza with antiviral agents. *Am J Epidemiol*, 159(7):623–633, 2004.

[28] Y. Yang, M.E. Halloran, and I.M. Longini Jr. A Bayesian model for evaluating influenza antiviral efficacy in household studies with asymptomatic infections. *Biostatistics*, 10(2):390, 2009.

[29] Y. Tang, I. Longini, and M. Halloran. Design and evaluation of prophylactic interventions using disease incidence data from close contact groups. *Applied Statistics*, 55(3):317–330, 2006.

[30] D.O. Scharfstein, M.E. Halloran, H. Chu, and M.J. Daniels. On estimation of vaccine efficacy using validation samples with selection bias. *Biostatistics*, 7(4):615, 2006.

[31] R. Glass, W. Beyeler, and H. Min. Targeted social distancing design for pandemic influenza. *Emerging Infectious Diseases*, 12(11):1671–1681, 2006.

[32] M. Lipsitch, T. Cohen, M. Murray, and B.R. Levin. Antiviral resistance and the control of pandemic influenza. *PLoS Medicine*, 4(1):111–115, 2007.

[33] Karima R. Nigmatulina and Richard C. Larson. Living with influenza: Impacts of government imposed and voluntarily selected interventions. *European Journal of Operational Research*, 195(2):613–627, 6/1 2009.

[34] I. M. Longini and J. S. Koopman. Household and community transmission parameters from final distributions of infections in households. *Biometrics*, 38(115-126), 1982.

[35] I. M. Longini. The generalized discrete-time epidemic model with immunity: a synthesis. *Math Biosci*, 82: 19–41, 1986.

[36] F. Ball and O. Lyne. Optimal vaccination policis for stochastic epdidemics among a population of households. *Mathematical Biosciences*, 177-178:333–354, 2002.

[37] C. Fraser, S. Riley, R. Anderson, N. Ferguson, and R. May. Factors that make an infectious disease outbreak controllable. *PNAS*, 101(16):6146–6151, 2004.

[38] F. Carrat, A. Lavenu, S. Cauchemez, and S. Deleguer. Repeated influenza vaccination of healthy children and adults:borrow now, pay later? *Epidemiol. Infect.*, 134:6370, 2005.

[39] J. Arino, F. Brauer, P.van den Driessche, J. Watmough, and J. Wu. Simple models for containment of a pandemic. *J. R. Soc. Interface*, 3:453–457, 2006.

[40] C. Dargatz, V. Georgescu, and L. Held. Stochastic modelling of the spatial spread of influenza in germany. *Austrian Journal of Statistics*, 35(1):520, 2006.

[41] N. Ferguson, S. Mallett, H. Jackson, N. Roberts, and P. Ward. A population-dynamic model for evaluating the potential spread of drug-resistant influenza virus infections during community-based use of antivirals. *Journal of Antimicrobial Chemotherapy*, 51:977–990, 2003.

[42] J. T. Wu, S. Riley, C. Fraser, and G. Leung. Spatial considerations for the allocation of pre-pandemic influenza vaccination in the United States. *Proc. R. Soc. Bio. Sci.*, 274(1627):2811–2817, 2006.

[43] J.K. Kelso, G.J. Milne, and H. Kelly. Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza. *BMC Public Health*, 9(1), 2009.

[44] N. Halder, J.K. Kelso, and G.J. Milne. Analysis of the effectiveness of interventions used during the 2009 A/H1N1 influenza pandemic. *BioMed Central*, 10(168), 2010.

[45] H. Yasuda and K. Suzuki. Measures against transmission of pandemic H1N1 influenza in Japan in 2009: simulation model. *Euro surveillance: bulletin européen sur les maladies transmissibles= European communicable disease bulletin*, 14(44), 2009.

[46] G.J. Milne, J.K. Kelso, H.A. Kelly, S.T. Huband, and J. McVernon. A small community model for the transmission of infectious diseases: comparison of school closure as an intervention in individual-based models of an influenza pandemic. *PLoS One*, 3(12), 2008.

[47] T. Das, and A. Savachkin. A large scale simulation model for assessment of societal risk and development of dynamic mitigation strategies. *IIE Transactions*, 40(9):893–905, 2008.

[48] N. M. Ferguson, D. Cummings, S. Cauchemez, C. Fraser, S. Riley, M. Aronrag, S. Lamsirithaworn, and D. Burke. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, 437: 209–214, 2005.

[49] T. Germann, K. Kadau, I. M. Longini, and C. Macken. Mitigation strategies for pandemic influenza in the United States. *PNAS*, 103:5935–5940, 2006.

[50] R. Patel, I. Longini, and M. Halloran. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology*, 234:201–212, 2005.

[51] A. Savachkin, A. Uribe, T. Das, and D. Prieto. Developing dynamic predictive strategies for mitigation of cross-regional pandemic outbreaks. *IIE Transactions (in review)*, 2010.

[52] A. Uribe, D. Prieto, A. Savachkin, T. Das, and Y. Zhu. A cross-regional pandemic outbreak simulation model: An aid to national resource allocation policy making. *Proceedings of the 3rd INFORMS Workshop on Data Mining and Health Informatics (DM-DI 2008)*, 2008.

[53] C. Mills, J. Robins, and M. Lipsitch. Transmissibility of 1918 pandemic influenza. *Nature*, 432:904–906, 2004.

[54] M. E. Halloran, N. M. Ferguson, S. Eubank, and I. Longini. Modeling targeted layered containment of an influenza pandemic in the United States. *PNAS*, 105(12):4639–4644, 2008.

[55] I. M. Longini, A. Nizam, X. Shufu, K. Ungchusak, W. Hanshaoworakul, D. Cummings, and M. E. Halloran. Containing pandemic influenza at the source. *Science*, 309:1083–1087, 2005.

[56] S. Eubank, H. Guclu, V. Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N.Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429:180–184, 2004.

[57] Committee on Modeling Community Containment for Pandemic Influenza. Modeling community containment for pandemic influenza: A letter report. `http://www.nap.edu/catalog/11800.html`, 2006. Last accessed on 12/09/2008.

[58] Centers for Disease Control and Prevention (CDC). Interim pre-pandemic planning guidance: Community strategy for pandemic influenza mitigation in the United States. `http://www.pandemicflu.gov/plan/community/community_mitigation.pdf`, 2007. Last accessed on 04/01/2009.

[59] HHS. Pandemic planning update iv, 2007. http://www.pandemicflu.gov/plan/panflureport4.html.

[60] J.F. Lawless and JF Lawless. *Statistical models and methods for lifetime data*. Wiley New York, 1982.

[61] O. Diekmann and JAP Heesterbeek. *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation*. Wiley, 2000.

[62] M. Bootsma and N. Ferguson. The effect of public health measures on the 1918 influenza pandemic in the us cities. *PNAS*, 104(18):7588–7593, 2007.

[63] J.F. Brundage and G.D. Shanks. Deaths from bacterial pneumonia during 1918–19 influenza pandemic. *Emerging Infectious Diseases*, 14(8):1193, 2008.

[64] Centers for Disease Control and Prevention (CDC). CDC influenza operational plan. `http://www.cdc.gov/flu/pandemic/cdcplan.htm`, 2006. Last accessed on 03/27/2009.

[65] S. Pasteur. Influenza A(H1N1) 2009 Monovalent Vaccine. `http://www.fda.gov/downloads/biologicsbloodvaccines/vaccines/approvedproducts/ucm182404.pfd`, 2009. Last accessed on 11/18/2009.

[66] R.J. Blendon, L.M. Koonin, J.M. Benson, M.S. Cetron, W.E. Pollard, E.W. Mitchell, K.J. Weldon, and M.J. Herrmann. Public response to community mitigation measures for pandemic influenza. *Emerging Infectious Diseases*, 14(5):778, 2008.

[67] M.Z. Sadique, W.J. Edmunds, R.D. Smith, W.J. Meerding, O. de Zwart, J. Brug, and P. Beutels. Precautionary behavior in response to perceived threat of pandemic influenza. *Emerging Infectious Diseases*, 13 (9):1307, 2007.

[68] R. Maunder, J. Hunter, L. Vincent, J. Bennett, N. Peladeau, M. Leszcz, J. Sadavoy, L. Verhaeghe, R. Steinberg, and T. Mazzulli. The immediate psychological and occupational impact of the 2003 SARS outbreak in a teaching hospital. *Canadian Medical Association Journal*, 168(10):1245–1251, 2003.

[69] E. Robertson, K. Hershenfield, S.L. Grace, and D.E. Stewart. The psychosocial effects of being quarantined following exposure to SARS: a qualitative study of Toronto health care workers. *Canadian Journal of Psychiatry*, 49:403–407, 2004.

[70] M.L. Pearson, C.B. Bridges, and S.A. Harper. Influenza vaccination of health-care personnel. *Recommendations of the Healthcare Infection Control Practices Advisory Committee (HICPAC) and the Advisory Committee on Immunization Practices (ACIP)[Internet]*, 2006. Last accessed on 10/28/2009.

[71] M. Keane, M. Walter, B. Patel, S. Moorthy, R. Stevens, K. Bradley, J. Buford, E. Anderson, L. Anderson, and K. Tibbals. Confidence in vaccination: a parent model. *Vaccine*, 23(19):2486–2493, 2005.

[72] V.P. Niederhauser, G. Baruffi, and R. Heck. Parental decision-making for the varicella vaccine. *Journal of Pediatric Health Care*, 15(5):236–243, 2001.

[73] S.D. Rhodes and K.C. Hergenrather. Exploring hepatitis B vaccination acceptance among young men who have sex with men: facilitators and barriers. *Preventive Medicine*, 35(2):128–134, 2002.

[74] S.L. Rosenthal, R.K. Kottenhahn, F.M. Biro, and P.A. Succop. Hepatitis B vaccine acceptance among adolescents and their parents. *Journal of Adolescent Health*, 17(4):248–255, 1995.

[75] M. Smailbegovic, G. Laing, and H. Bedford. Why do parents decide against immunization? The effect of health beliefs and health professionals. *Child: Care, Health & Development*, 29(4):303, 2003.

[76] Colardo Department of Human Services Division of Mental Health. Pandemic influenza: Quarantine, isolation and social distancing. `http://www.flu.gov/news/colorado_toolbox.pdf`, 2009. Last accessed on 10/28/2009.

[77] T.J. Safranek, D.N. Lawrence, L.T. Kuriand, D.H. Culver, W.C. Wiederholt, N.S. Hayner, M.T. Osterholm, P. O'Brien, and J.M. Hughes. Reassessment of the association between Guillain-Barré syndrome and receipt of swine influenza vaccine in 1976-1977: results of a two-state study. *American journal of epidemiology*, 133(9):940, 1991.

[78] K.M. Cummings, A.M. Jette, B.M. Brock, and D.P. Haefner. Psychosocial determinants of immunization behavior in a swine influenza campaign. *Medical Care*, 17(6):639–649, 1979.

[79] The New Yorker. The fear factor. `http://www.newyorker.com/talk/comment/2009/10/12/091012taco_talk_specter`, 2009. Last accessed on 10/28/2009.

[80] The New York Times. Doctors swamped by swine flu vaccine fears. `http://www.msnbc.msn.com/id/33179695/ns/health-swine_flu/`, 2009. Last accessed on 10/28/2009.

[81] U.S Census Bureau. 2001 American community survey. `http://www.census.gov/prod/2001pubs/statab/sec01.pdf`, 2000. Last accessed on 03/27/2009.

[82] Bureau of transportation statistics. 2001 National household travel survey (NTHS). http://www.bts.gov/programs/national_household_travel_survey/, 2002. Last accessed on 12/09/2008.

[83] A. Savachkin, A. Uribe-Sánchez, T. Das, D. Prieto, A. Santana, D. Martinez. Supplemental data and model parameter values for cross-regional simulation-based optimization testbed. http://imse.eng.usf.edu/pandemic/supplement.pdf, 2010. Last accessed on 04/15/2010.

[84] Tampa International Airport: daily traffic volume data. http://www.tampaairport.com, 2010. Last accessed on 04/07/2010.

[85] Miami International Airport: daily traffic volume data. http://www.miami-airport.com, 2010. Last accessed on 04/07/2010.

[86] Jacksonville Aviation Authority: daily traffic volume data. http://www.jaa.aero/General/Default.aspx, 2010. Last accessed on 04/07/2010.

[87] Tallahassee Regional Airport: daily traffic volume data. http://www.talgov.com/airport/index.cfm, 2010. Last accessed on 04/07/2010.

[88] B. Ortutay. How thermal-imaging cameras spot flu fevers. http://www.msnbc.msn.com/id/30523865/, 2010. Last accessed on 03/19/2010.

[89] Writing committee of the World Health Organization (WHO). Consultation on human influenza A/H5. Avian influenza A(H5N1) infection in humans. *N Engl J Med*, 353(13):1374–1385, 2005.

[90] M. Meltzer, N. Cox, and K. Fukuda. The economic impact of pandemic influenza in the United States: Priorities for intervention. *Emerging Infectious Deseases*, 5(5):659–671, 1999.

[91] World Health Organization (WHO). WHO guidelines on the use of vaccine and antivirals during influenza pandemics. http://www.who.int/csr/resources/publications/influenza/11_29_01_A.pdf, 2004. Retreived 03/27/2009.

[92] Institute of Medicine (IOM). Antivirals for pandemic influenza: Guidance on developing a distribution and dispensing program. *The National Academies Press*, 2008.

[93] J. Treanor, J. Campbell, K. Zangwill, T. Rowe, and M. Wolff. Safety and immunogenicity of an inactivated subvirion influenza A(H5N1) vaccine. *N Engl J Med*, 3554(13):1343, 2006.

[94] PayScale. Job: Registered nurse (rn). http://www.payscale.com/rcsearch.aspx?country=US&str=Registered+Nurse+, 2009. Last accessed on 12/28/2009.

[95] Centers for Disease Control and Prevention. CDC vaccine price list. http://www.cdc.gov/vaccines/programs/vfc/cdc-vac-price-list.htm#flu, 2009. Last accessed on 12/28/2009.

[96] PharmacyChecker.com. Pricing & ordering comparisons. http://www.pharmacychecker.com/Pricing.asp?DrugId=36260&DrugStrengthId=61300&sortby=Price, 2009. Last accessed on 12/28/2009.

[97] R.J. Blendon, C.M. DesRoches, M.S. Cetron, J.M. Benson, T. Meinhardt, and W. Pollard. Attitudes toward the use of quarantine in a public health emergency in four countries. *Health Affairs*, 25(2):15–25, 2006.

[98] A. Svensson. A note on generation times in epidemic models. *Math Biosci*, 208(1):300–311, 2007.

[99] T. Halfhill. Inflation calculator. `http://www.halfhill.com/inflatation.html`, 2009. Last accessed on 12/04/2009.

[100] V. Colizza, A. Barrat, A. Valleron M. Barthelemy, and A. Vespignani. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLOS Medicine*, 4:95–110, 2007.

[101] P. Cooley, L. Ganapathi, G. Ghneim, S. Holmberg, W. Wheaton, and C. Hollingsworth. Using influenza-like illness data to reconstruct an influenza outbreak. *Mathematical and Computer Modelling*, (48):929–939, 2008.

[102] Foundation Eclipse. The Spatio-Temporal Epidemiological Modeler. `http://www.eclipse.org/stem/intro.php`, 2009. Last accessed on 11/06/2009.

ABOUT THE AUTHOR

Andrés Uribe-Sánchez received his B.S. in Industrial Engineering with emphasis on Operations Research and Finance from the University of Los Andes, Bogota, Colombia in 2003. In 2006, he received his M.S. in Management Systems at the University of Puerto Rico at Mayagez. He received his Ph.D. in Industrial Engineering from the University of South Florida in 2010. In April 2010, he was awarded the Distinguished Graduate Achievement Award from the University of South Florida. His areas of research interest include engineering risk analysis, support of healthcare enterprise capacity management, and decision support for mitigation of large-scale public health disasters.