

Improving Product Availability in Hospitals: the Role of Inventory Inaccuracies

by
David C. Opolon

B.S., Science and Executive Engineering, MINES ParisTech, France, 2003
M.Sc., Science and Executive Engineering, MINES ParisTech, France, 2005

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January 2010

© 2010, Massachusetts Institute of Technology. All rights reserved.

The author hereby grants to MIT the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author _____

David C. Opolon
Engineering Systems Division

Certified by _____

Yossi Sheffi
Elisha Gray II Professor of Engineering Systems, Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by _____

Stephen C. Graves
Abraham J. Siegel Professor of Management Science,
Professor of Mechanical Engineering and Engineering Systems

Certified by _____

Kimberly M. Thompson
Associate Professor of Risk Analysis and Decision Science, Harvard School of Public Health

Accepted by _____

Nancy Leveson
Professor of Aeronautics and Astronautics and Engineering Systems
Chair, Engineering Systems Division Education Committee

[Page intentionally left blank]

Improving Product Availability in Hospitals: the Role of Inventory Inaccuracies

by

David C. Opolon

Submitted to the Engineering Systems Division on January 31, 2010 in partial fulfillment of
the requirements for the Degree of Doctor in Philosophy in Engineering Systems

Abstract

All players in the healthcare industry face increasing public and political pressure to improve quality of care and control costs. Hospitals, on the frontline of this challenge, face nursing shortages and financial constraints. Survey data indicate that missing medication and supplies interrupt nurses more than twice per shift, increasing costs and putting patients at risk. These challenges persist even though over 72% of U.S. hospitals have deployed Automated Dispensing Machines (ADMs), electronic cabinets that automate inventory management processes and improve product availability.

This research investigates the role of inventory inaccuracies, i.e., mismatches between book inventory and physical inventory on hand, as drivers of product availability in hospitals. The research objectives are three-fold: (1) characterize the sources of inventory inaccuracies prevalent in a hospital context; (2) quantify the impact of inventory inaccuracies on product availability and performance metrics; and (3) identify and evaluate practical strategies that hospitals can use to improve product availability by reducing and mitigating inventory inaccuracies.

This thesis views the hospital supply chain as a socio-technical system and addresses the research questions using a multilevel, multi-method approach. The research is empirically grounded by the case study of Lambda, a New England area hospital that provided qualitative and high-frequency transactional data from its network of 108 ADMs that stock over 21,000 product-location combinations.

First, by classifying sources of inventory inaccuracies this thesis identifies Imperfect Demand Recording as a hospital-specific source of such inaccuracies. Recording Accuracy is proposed as a metric of user behavior at product and location levels, and reveals that between five and thirty percent of product usage is not recorded. Then, a single-product Discrete-Event Simulation (DES) model shows that Imperfect Demand Recording causes large reductions in availability unless mitigated by frequent and consistent (i.e., equally-spaced) inventory counts, and that service level estimates provided by ADMs can have a large, optimistic bias. Assuming that count timing is independent of inventory state, an analytical model provides a closed-form generalization of the simulation results and shows that variability in cycle count has a nonlinear and substantial effect, causing 35% of counts performed at Lambda to be ineffective. Finally, a sequential and iterative framework integrating the managerial implications of these contributions is proposed.

Doctoral Committee:

Yossi Sheffi (Thesis Supervisor)

Title: Elisha Gray II Professor of Engineering Systems,
Professor of Civil and Environmental Engineering

Stephen C. Graves

Title: Abraham J. Siegel Professor of Management Science,
Professor of Mechanical Engineering and Engineering Systems

Kimberly M. Thompson

Title: Associate Professor of Risk Analysis and Decision Science,
Harvard School of Public Health

Acknowledgments

First, my biggest thanks go to my advisor Prof. Yossi Sheffi, for giving me the opportunity to work with and learn from him. You took a chance when accepting an unproven 20-year old as a visiting student in the Supply Chain Response to Terrorism group, and made me discover how much I enjoyed research. You were always supportive of my academic interests and magically balanced encouragement and constructive criticism. I could not have envisioned a better mentor for this journey.

I would also like to deeply thank my committee members, Prof. Steve Graves and Prof. Kimberly Thompson. Steve, I thank you for generously making yourself available to share your invaluable insights on inventory management, providing guidance on models and assumptions, and giving me the honor and wonderful experience of assisting you in teaching Supply Chain Management to Sloan MBAs. Kim, thank you for pushing me to see the hospital through the eyes of clinicians and always suggesting ways to make this research more impactful for them and at the same time more academically rigorous.

This research would not have been possible without the people of Lambda hospital. Sandy and Sue, you opened the door of your organization to CTL and gave me an unbelievable level of access. John, you were always willing to share your knowledge of the inventory processes and challenges faced by materials managers, and answer my emails and welcome me to the warehouse during the busy early morning work. Beth and Marty, you worked hard to help me get the data I needed. I hope that this research has been helpful and see many more opportunities to apply it going forward.

Kathryn Schumacher played a pivotal role in this research by helping implement and exercise the simulation model. Our conversations led to many of the core insights of this thesis. I am extremely fortunate to have found an UROP as dedicated, capable and thoughtful as you, and ultimately, a friend.

Thank you to those who provided feedback and advice during the course of this research. In particular, I would like to thank Prof. Gabriel Bitran for suggesting that I investigate the role of variability in cycle counting further, after seeing the results of the simulation at the Operations Management & System Dynamics seminar. Thank you to Dr. Philippe Bonnefoy, Romain Lévy and Esya Volchek for helping me prepare my defense and being supportive friends throughout this process.

The members of the MIT Efficient Healthcare Delivery (MEHD) group, and in particular Dr. Mahender Singh, shaped my understanding of the health care supply chain. Mahender, thank you for your feedback and support on the early stages of this research. I feel very privileged to have benefited from your creativity in research. Caroline, Pierre-Eloi, Shardul, Stéphane: thank you for your help, support, and friendship. Thank you to the industry sponsors of MEHD and the MIT Center for Transportation and Logistics for supporting this research financially and generously sharing your industry knowledge.

My journey at MIT started in 2003 and I would like to thank those who made it such a wonderful experience. Prof. Richard de Neufville – thank you for your ability to connect with students around the world in their native language and make them feel at home. Prof. John Sterman – thank you for teaching me System Dynamics, and thus changing my mental model of the world works. Jim Rice – thank you for helping me get around MIT and your continued encouragements during the difficult and the good times of the Ph.D. Dr. Shalom Saar – thank you for making me more self-aware, and providing advice when I most needed it.

To all my ESD colleagues, and in particular: Blandine Antoine, David Broniatowski, Erica Gralla, Joao Castro, Tony Craig, Sid Rupani, Stephan Feilhauer, Roberto Perez-Franco, Valerie Karplus – thank you for your camaraderie, friendship and willingness to hear my research many times over as it developed!

My friends, both in France and at MIT, have been terrific: Ali Motamedi, Anne-Laure de Chammard, Carine Simon, Christophe Antoine, David Gadmer, David Wang, Durga Pandey, Emilie Bonnefoy, Francois de Nanteuil, Gregoire de Chammard, Guillaume de Roo, Hampus Hillerstrom, Kai Trepte, Jonathan Kluberg, Larry Chang, Laurent Guerin, Leeland Ekstrom, Lorena Lee-Houghton, Louis Mosley, Lucille Guillaud, Manolis Kellis, Michael Mestre, Michel Dahan, Macha Belkova, Nan Gu, Nicholas Veg, Noémie Chocat, Oaz Nir, Pankaj Sarin, Paul Monasterio, Philipp Robbel, Raphael Berger, Ray Rahman, Reynaldo Sandoval, Romain Lacombe, Stanislas de Margerie, Tanguy Chau, Vincent Cordoliani, Yoda Patta, Ziad Sultan – I am blessed to count you as my friends.

Finally, this thesis is dedicated to my family. To my parents, Paule and Pierre, and my siblings, Nathalie and Claude, for always believing in me and supporting me throughout this journey. To my love, Coppélia, for feeding me exquisite food, editing this thesis, embodying focus and creativity, being a constant source of inspiration and providing joy and love in my life!

Contents

Front Matter

Contents 7

List of Tables..... 12

List of Figures..... 15

List of Appendices..... 18

1 Introduction 19

1.1 Background and Motivation 19

1.2 Research Opportunity: the Hospital Internal Supply Chain as an Engineering System 20

1.3 Research Objectives..... 21

 1.3.1 Characterizing the sources of inventory inaccuracies 22

 1.3.2 Impact of inventory inaccuracies on system performance..... 22

 1.3.3 Practical strategies for the reduction and management of inventory inaccuracies in the hospital 23

1.4 Dissertation Structure and Overview 24

2 The Hospital Supply Chain 27

2.1 The Hospital Supply Chain within the Health Care Delivery System 27

 2.1.1 Overview of the U.S. health care delivery system 28

 2.1.2 The hospital within the health care supply chain 29

2.2	The Hospital Supply Chain as a Socio-Economic-Technical System.....	31
2.2.1	Objectives	33
2.2.2	Stakeholders	35
2.2.3	Technical infrastructure.....	41
2.2.4	Product classification.....	42
3	Related Work	45
3.1	Hospital Inventory Management.....	45
3.1.1	A note on hospital supply chain management	45
3.1.2	Hospital inventory management	46
3.1.3	RFID applications.....	48
3.2	Inventory Inaccuracy.....	48
3.2.1	Empirical studies of Inventory Inaccuracy	49
3.2.2	Effect of Inventory Inaccuracy on service levels	51
3.2.3	Counting and ordering policies under Inventory Inaccuracy.....	54
4	Problem Statement and Approach	57
4.1	Problem Statement.....	57
4.1.1	Discussion of the literature	58
4.1.2	Research questions	59
4.2	Research Approach.....	60
4.2.1	Research site and scope.....	61
4.2.2	Automated Dispensing Machine field data	61
5	Characterizing and Quantifying Sources of Inventory Inaccuracies	65
5.1	Typology of Sources of Inventory Inaccuracies	65
5.1.1	Imperfect Demand Recording	66

5.1.2	Incorrect returns.....	67
5.1.3	Incorrect replenishments	68
5.1.4	Incorrect counts	69
5.1.5	Fake transactions	69
5.2	Existing Metrics of Inventory Inaccuracy	70
5.2.1	Inaccuracy detection.....	70
5.2.2	Existing reports	71
5.2.3	Limitations of Existing Metrics	72
5.3	The Role of Framing.....	76
5.4	Quantifying Imperfect Demand Recording	77
5.4.1	Challenges to estimating Recording Accuracy from ADM data.....	78
5.4.2	Estimation methodology for a single product	79
5.4.3	Estimation for multiple products and with heterogeneity	84
5.5	Empirical Analysis of Recording Accuracy	87
5.5.1	Data pre-processing.....	87
5.5.2	Implementation	88
5.5.3	Product groups.....	89
5.5.4	Results.....	90
6	Impact of Inventory Inaccuracies on System Performance	95
6.1	Discrete-Event Simulation Model	95
6.1.1	Key assumptions.....	95
6.1.2	Model structure.....	97
6.1.3	Implementation	100
6.1.4	Base case	101

6.2	Effect of Imperfect Demand Recording on Service Level Metrics.....	103
6.2.1	Fill rate and physical in-stock probability	105
6.2.2	Physical stock-out frequency.....	108
6.2.3	Type I service level	110
6.2.4	Sensitivity to variability in Count Periodicity	112
6.3	Investigating the Bias in Book View Metrics	114
6.3.1	Bias in the book view in-stock probability	114
6.3.2	Bias in the book view stock-out frequency	116
6.3.3	Counting on replenishment policy	120
6.4	Evaluation of Strategies to Improve Product Availability.....	122
6.4.1	Scenario definition and performance metrics	123
6.4.2	Sample panel definition.....	126
6.4.3	Simulation results	129
6.5	Summary.....	135
7	Variability in Cycle Counting	137
7.1	Service Level Optimality of Equally Spaced Counts.....	139
7.1.1	Assumptions and notations.....	139
7.1.2	Derivation	142
7.2	Quantifying the Service Level Deterioration Resulting From Variability.	148
7.2.1	Derivation	148
7.2.2	Count Inefficiency metric	150
7.2.3	Numerical study	152
7.3	Empirical Study of Intercount Times.....	157
7.3.1	Measuring Intercount Time Variability	157

7.3.2	The case of the Cardiac Catheterization lab	160
7.3.3	Testing for dynamic counting strategies	163
7.3.4	Estimating the count reduction potential at Lambda	165
8	Managerial Implications	169
8.1	Evaluation of Inventory Inaccuracy Strategies	169
8.1.1	Reducing Imperfect Demand Recording.....	169
8.1.2	Counting more frequently and reducing Count Variability	171
8.1.3	Counting on Replenishment.....	172
8.1.4	Dynamic Inventory Policy	173
8.2	A Holistic and Iterative Framework for Improving Product Availability in the Hospital.....	174
8.2.1	Framework description	174
8.2.2	Discussion	177
9	Conclusions & Contributions	179
9.1	Conclusions	179
9.1.1	Characterization of Inventory Inaccuracies in the Hospital.....	179
9.1.2	Drivers of Product Availability	180
9.1.3	Bias and Inadequacy of Existing Service Level Metrics.....	180
9.1.4	Implications for Hospital Inventory Management	181
9.2	Contributions	182
9.3	Suggestions for Future Research.....	183
10	Bibliography	184

List of Tables

Table 3.1: Inventory Inaccuracy definitions and magnitude in empirical studies	50
Table 3.2: Summary of the characteristics of models of the service level impact of inventory inaccuracy	53
Table 4.1: Sample Transaction Record for Item 61367 in Station Cath Lab A.....	63
Table 5.1: Discrepancy detection rates for different transaction types.....	71
Table 5.2 Distribution of negative values among count events	88
Table 6.1: Summary of input parameters and base case values.....	103
Table 6.2: Fill Rate and Physical In-stock Probability for different values of the Recording Accuracy and Count Periodicity	107
Table 6.3: Physical View Stock-out Frequency (# stock-outs / week) for different values of the Recording Accuracy and Count Periodicity.	109
Table 6.4: Simulated Type I Service Level (Probability of stock-out in a replenishment cycle) for different values of the Recording Accuracy and Count Periodicity, as well as predicted Type I Service Levels according to Morey (1985) model.....	111
Table 6.5: Physical and Book View In-stock Probability for different values of the Recording Accuracy and Count Periodicity	115
Table 6.6: Physical and Book View Stock-out Frequency (# stock-outs / week) for different values of the Recording Accuracy and Count Periodicity	117

Table 6.7: Scenario characteristics. All combinations of the scenarios above were simulated for each item in the panel, corresponding to 84 total inventory inaccuracy mitigation policies.....	123
Table 6.8: Summary of Station-SKU Sample Panel (CD: Closed Door, OS: Open Shelf).....	128
Table 6.9: Summary of the bias in different product availability metrics.	136
Table 7.1: Service Level Breakdown for different values of α , the target service level in the absence of inventory inaccuracies ($K = 0.5$).....	156
Table 7.2: Average Coefficient of Variation of inter-count times broken down by the number of cycle counts performed during the 15-month study period	158
Table 7.3: Summary statistics on the distribution of variability of inter-count times across different products in a station (one station-SKU = one observation)	160
Table 7.4: Illustrative Inter-count Time CV and Inter-count Demand CV for different Count Days, taking into account nonstationarity of demand during the week. This table assumes that counts are executed before the work day on those dates.	162
Table 7.5: Proportion of products for which the null hypothesis is rejected at the 0.05 level (stations with less than 50 SKUs were excluded).....	167
Table 9.1: Service Level Breakdown for different values of K , ($\alpha = 0.80$)	246
Table 9.2: Service Level Breakdown for different values of K , ($\alpha = 0.90$)	247
Table 9.3: Service Level Breakdown for different values of K , ($\alpha = 0.95$)	248
Table 9.4: Service Level Breakdown for different values of K , ($\alpha = 0.98$)	249
Table 9.5: Service Level Breakdown for different values of K , ($\alpha = 0.99$)	250
Table 9.6: Average of Mean Inter-Count Time, Average of C.V., Number of SKUs, and Number of SKUs for which we reject the null at the 0.01, 0.05 and 0.10 significance levels	254

List of Figures

Figure 2-1: Flows of Money and Services in a Health Care Delivery system (Berndt, 2007)	28
Figure 2-2: The Health Care Supply Chain (Oliveira, 2005).....	30
Figure 2-3: Percentage of hospital CEOs citing issue among three most important (ACHE, 2009)	34
Figure 2-4: Projected full-time equivalent nurses demand and supply (Biviano, et al., 2007)	36
Figure 4-1: Conceptual view of the Social, Technical and Managerial components of the Hospital Supply Chain as an Engineering System	58
Figure 4-2: Diagram summarizing the multi-method approach to the research questions	61
Figure 5-1 A “For patients’ sake, Press Take” sticker affixed to an Automated Dispensing Machine at Lambda hospital	67
Figure 5-2: Sample excerpt of Null Transactions report (user name obfuscated)	73
Figure 5-3: Sample Compliance Report	74
Figure 5-4: Internal slide used at Lambda during materials management technicians training, illustrating the sensibility of Quantity Compliance (QC) to the number of counts and outlier observations	75
Figure 5-5 Count events (blue dots), periods and corresponding allocation weights	83
Figure 5-6 Hierarchical representation of the Beta-Binomial Model (BBM).....	85

Figure 5-7 Beta-Binomial Model: Monthly estimates of the Recording Accuracy for seven stations	91
Figure 5-8: Histogram of Monthly Recording Accuracy, separating Lambda's Main facility (left pane) from Lambda North (right pane).....	92
Figure 5-9: Stations ranked by Median Recording Accuracy across 15 months.....	93
Figure 6-1: Key processes used in the simulation. The clock signifies periodicity.....	98
Figure 6-2: Effect of Count Periodicity on In-Stock Probability.....	106
Figure 6-3: Effect of Variability of Inter-Count Times on In-Stock Probability, with a count periodicity of 21 days, a truncated normal distribution (C.V. = 0.5) and an exponential distribution (C.V. = 1).....	113
Figure 6-4: Physical and Book Stock-out Frequencies as a function of Recording Accuracy, at a Count Periodicity of 14 days.....	116
Figure 6-5: Stock-Out Detection Rate as a function of Recording Accuracy, at different Count Periodicities.....	118
Figure 6-6: Physical and Book Stock-out Frequencies under two scenarios: no counting on replenishment and always counting on replenishment.	122
Figure 6-7: Efficiency frontier for item 63795: Disposable Bed Pan.	131
Figure 6-8: Efficiency frontier for item 63795: Disposable Bed Pan. Reduced count variability scenarios are depicted with full symbols.	132
Figure 6-9: Efficiency frontier for item 56840: Stone Tipless Extractor. Reduced count variability scenarios are depicted with full symbols.....	133
Figure 6-10: Efficiency frontier for item 63795: Disposable Bed Pan. Reduced count variability scenarios are depicted with full symbols.	135
Figure 7-1: Simplified timeline of count events	146

Figure 7-2: Count Inefficiency as a function of Count Variability, showing an S-curve pattern 151

Figure 7-3: Service Level as a function of the Mean Time between Counts..... 153

Figure 7-4: Breakdown of Service Level into its different values as a function of Count Variability ($K = 0.5$, $\alpha = 0.95$). As inter-count time variability increases, the service level impact of inequally spaced counts increases and the bound is less tight..... 155

Figure 7-5: Histogram of the Coefficient of Variation (CV) of inter-count times for 6,110 hospital Station-SKUs..... 159

Figure 7-6: Total number of Count, Demand and Replenishment transactions by day of the week over the 15-month study period for Station A of the Cardiac Catheterization Lab..... 161

Figure 8-1: Causal Loop Diagram of the Iterative Improvement Framework The double slash indicates the delay between Imperfect Demand Recording and the Recording Accuracy estimate. 176

List of Appendices

Appendix A. Interview list	194
Appendix B. Investigation of Demand Autocorrelation	195
Appendix C. Analytical Model of Demand Recording	198
Appendix D. Discrete-Event Simulation Model Code	205
Appendix E. Study of p for Two-Sided Inaccuracies	238
Appendix F. Study of the Convexity of MAPL	242
Appendix G. Sensitivity Analysis for Parameter K	246
Appendix H. Count Variability in Hospital Stations	252
Appendix I. Sample Panel Simulation Results	255

Chapter 1

Introduction

1.1 Background and Motivation

U.S. health care costs are the highest in the world (OECD, 2009). In 2007, expenditures on health care reached \$7,421 per capita, amounted to 16.2% of GDP and kept increasing faster than inflation (CMS, 2009). This represents a tremendous burden for all payers: households whose health insurance premiums increased at an average rate of 10% per year between 1999 and 2006, corporations whose competitiveness is diminished by employee health care costs, and the government whose budget deficit (and projected future deficits) is due, in part, to the Medicare and Medicaid programs.

Beyond costs, according to some observers, the U.S. health care system is not delivering an acceptable level of performance (Reid, 2005). In 2005, the National Academy of Engineering and Institute of Medicine published a joint report entitled “Building a Better Delivery System, a New Engineering/Healthcare Partnership” recommending the application of systems engineering tools to promote the following objectives for the health care system: Safety, Effectiveness (avoiding under use and overuse), Patient-centric care, Timeliness, Efficiency and Equity. While some local improvement efforts have been achieved using systems engineering tools, the report notes that information technology-intensive enterprise management and supply chain management have not been applied strategically to measure and improve performance (Reid, 2005).

Within the health care industry, hospitals face strong financial and operational pressures. In 2004, over 25% of hospitals had a negative margin and the aggregate operating margin for the industry was 3.6% (AHA & Levin Group, 2006). Operationally, and in spite of the 2008-2009 recession, hospitals faced difficulties hiring and retaining nurses as the overall nurse population ages (BLS, 2010). In an observational study of 26 nurses in nine U.S. hospitals, Tucker (2004) reports that time spent resolving operational failures accounted for 9% of total nursing time, and was almost equal to overtime. Problems originating outside the nursing unit, such as a late medication or a shortage of swabs necessary for taking a sample for infection testing, accounted for 55% of failures and had a median cost of \$124 per incident. In a subsequent survey of nurses in 48 U.S. hospitals, nurses reported an average of 4.5 operational failures per shift, with failures linked to supply items averaging 1.2 times per shift (Tucker & Spear, 2006). Therefore, insufficient product availability is a frequent problem in U.S. hospitals, and results in additional nursing and management time, delayed procedures and workflow interruptions that put patients at risk of medical errors (Tucker & Spear, 2006).

1.2 Research Opportunity: the Hospital Internal Supply Chain as an Engineering System

Given the societal and economic importance of the healthcare and hospital industries in developed nations and the burden of the issues faced by practitioners, it is surprising how little attention its supply chain has received to date, at least in comparison with manufacturing. Within the health care supply chain, hospitals seem to have received the least attention, despite the fact that this “last mile” suffers from significant challenges (Singh & Rice, 2006; Tucker, 2004; Tucker & Spear, 2006). For instance, inventory turns and labor utilization are typically lower in hospitals materials management departments than at distribution centers due to the economies of scale and scope. From the viewpoint of hospital executives, the most natural explanation is that supply chain management is often considered a peripheral or ancillary function - usually referred to as “materials management” - relative to clinical services that constitute the core mission of the organization (Singh & Rice, 2006). For this reason, until recently hospitals may have been a less natural research ground for supply chain management academics.

In 2005, 72% of U.S. hospitals had deployed Automated Dispensing Machines (ADMs), electronic cabinets that automate inventory management processes (Pedersen, Schneider, & Scheckelhoff, 2006). However, inventory inaccuracies, i.e. mismatches between the book inventory and physical inventory on hand, are common in these systems (Klibanov & Eckel, 2003). In a case study of Automated Dispensing Machines in a Midwestern hospital, DeScioli (2005) reports anecdotal evidence that demand is not always recorded by end-users, a phenomenon which this thesis refers to as Imperfect Demand Recording. Of course, inventory inaccuracies are not unique to hospitals and were first acknowledged as a significant problem in the 1960's, in particular by Rinehart (1960) within government supply centers. However, the advent of technologies enabling (or claiming to achieve) better inventory visibility, such as Radio-Frequency Identification (RFID), has prompted further research into inventory inaccuracy since 2003, which is reviewed in Chapter 3 of this thesis. Most of the recent literature focuses on retail environments and treats inventory inaccuracies as exogenous factors that are independent from demand.

1.3 Research Objectives

This thesis takes an Engineering Systems perspective, looking at the internal hospital supply chain, defined as starting at the dock of the hospital and ending at the patient bedside. It investigates the role of inventory inaccuracies in product availability.

The thesis has three research objectives: (1) characterize the sources of inventory inaccuracies prevalent in a hospital context; (2) quantify the impact of inventory inaccuracies on product availability and performance metrics; and (3) identify and evaluate practical strategies that hospitals can use to improve product availability by reducing and mitigating inventory inaccuracies.

The research questions are addressed using a multilevel (i.e. at the product, ward, and hospital levels) and multi method, systems approach, combining statistical methods, Discrete-Event Simulation (DES), and analytical models. The research is empirically grounded by the case study of Lambda, a New England area hospital that provided qualitative and high-frequency transactional data from its network of 108 ADMs that stock over 21,000 product-location combinations.

1.3.1 Characterizing the sources of inventory inaccuracies

The first step of this thesis is to qualitatively and quantitatively characterize the sources of inventory inaccuracies in terms of user behavior within the context of the operations of a hospital, and understand how these inaccuracies are mitigated.

- **(Q1a)** *Can we identify and quantify specific user behaviors that introduce inventory inaccuracies in the hospital?*
- **(Q1b)** *How are inventory inaccuracies currently mitigated in the hospital?*

Research question (Q1a) aims at establishing a qualitative and quantitative understanding of the sources of inventory inaccuracies within the context of a hospital's inventory management processes. First, this thesis provides a background assessment and allows us to evaluate assumptions about inventory inaccuracies made in recent inventory management models that account for inventory inaccuracies. Second, we develop a measurement methodology to assist in identifying problem areas, targeting improvement opportunities, and providing feedback to users in order to drive behavioral change. Research question (Q1b) focuses on existing counting policies and is addressed in section 7.3.

1.3.2 Impact of inventory inaccuracies on system performance

The second point investigated is the impact of inventory record inaccuracies on product availability and metrics of product availability. The specific research questions are:

- **(Q2a)** *What is the quantitative impact of inventory record inaccuracies on book and physical product availability?*
- **(Q2b)** *How do counting policies affect product availability?*

1.3.3 Practical strategies for the reduction and management of inventory inaccuracies in the hospital

Much of the research in the area of inventory inaccuracy is within the retail sector. Within this context, DeHoratius, Mersereau and Schrage (2008) identify “at least three ways a retailer may respond to inventory record inaccuracy:

1. “Prevention: Reduce or eliminate the root causes of inventory record inaccuracy through the implementation and execution of process improvement.”
2. “Correction: Identify and correct existing inventory record discrepancies through auditing policies.”
3. “Integration: Use inventory planning and decision tools robust enough to account for the presence of record inaccuracy.” (DeHoratius, et al., 2008, p. 258)

In addition to these three ways, hospitals sometimes respond by adding inventory to buffer the inventory record uncertainty (i.e., “safety stock”). However, this strategy is often limited by the lack of physical space in the hospitals (more specifically, by the opportunity cost of inventory storage space in the hospital relative to a revenue-generating room), by the opportunity cost of capital, and by the costs of wasted products due to obsolescence.

The benefits of prevention, correction and integration are assessed through sensitivity analyses on the quantitative models established in answering (Q2). Research question (Q3) can be summarized in the following way:

- **(Q3)** *What are effective strategies that hospital managers can use, recognizing the interdependencies between different approaches?*

As an example of such interdependencies, consider a Correction strategy consisting of more frequent physical audits. First, Correction purges the inventory inaccuracies in the system and mitigates their consequences. Second, by providing more data on such inventory inaccuracies, it al-

lows for improved measurement of the drivers of inventory inaccuracies (using the methodologies derived in Q1a), therefore supporting Prevention-focused process improvements efforts.

Deriving an “optimal” strategy would require extensive modeling of the (uncertain) costs involved in implementing all three approaches and their combinations. Therefore, the approach in (Q3) is not to seek such an “optimal” strategy, but rather to evaluate the effectiveness of different operationally-realistic scenarios, providing the manager with insights into the tradeoffs and expected effects of pursuing different strategies, depending on the current level of inaccuracies in the system.

1.4 Dissertation Structure and Overview

This section summarizes the purpose of each subsequent chapter.

Chapter 2 provides background on the structure of the hospital supply chain. A generic hospital supply chain is described as a socio-technical system in terms of stakeholders with different objectives, and placed within the context of the broader U.S. healthcare system.

Chapter 3 is a review of different streams of literature relevant to this thesis. First, research on hospital supply chain management is considered, with a particular focus on inventory management. Second, research on inventory inaccuracy is reviewed, including descriptive work assessing the prevalence of the problem in different industries, studies of the impact of inventory inaccuracies on system performance and potential mitigation strategies.

Chapter 4 expands the detailed research questions and describes the research approach used to address them. First, the characteristics of the hospital supply chain that have not been addressed in the literature, in particular underlying assumptions about the process generating inventory inaccuracies, and differences in the availability of frequent count data, are discussed. Lambda hospital, a 300-bed New England hospital that has invested in Automated Dispensing Machines (ADMs) is introduced as an in-depth case study of a hospital supply chain. As part of this case study, over 3 million transaction records, covering a period of over 15 months, were collected for data analysis. Even if the extent of inventory inaccuracies varies among hospitals, it is important

to note that Automated Dispensing Machines are commonly in use at many U.S. hospitals, and therefore the methodologies developed and analyzed in this case study are widely applicable.

Chapter 5 characterizes the sources of inventory inaccuracies at Lambda (Q1). First, potential sources of inventory inaccuracies are listed and their relative importance is assessed, based on data collected through hospital interviews and quantitative transactional data. Imperfect Demand Recording is identified as a major source of inventory inaccuracies at Lambda. The root causes of this phenomenon are discussed, a metric for characterizing behavioral sources of inventory inaccuracy is introduced, and inventory inaccuracies are quantified using Lambda's transaction data according to this quantitative methodology.

Chapter 6 describes a Discrete-Event Simulation model of inventory inaccuracy in the hospital. First, the model's assumptions, structure and output are described. Throughout the model, the true (or physical) state of the system, defined as the "Physical View" (PV) is distinguished from the state of the system as recorded in the information system, defined as the "Book View" (BV). This modeling paradigm has multiple advantages: (1) it enables modeling the effects of information inaccuracy; (2) it ensures that all operational decisions are executed using only the information available in the "Book View"; and (3) the analysis of the output data can compare and contrast both views, potentially highlighting areas where "Book View" metrics are biased by inventory inaccuracies, providing the methodology for addressing questions (Q2a), (Q2b). A possible bias between real and measured performance is identified and analyzed for different performance metrics under different inventory counting policies. Variability in the time between counts is identified as a factor that reduces product availability. Multiple strategies to improve product availability are evaluated on a sample of products from the hospital using a factorial simulation experiment.

Chapter 7 generalizes the observations made in chapter 6 on the effects of count variability using an analytical model, and provides empirical evidence of count variability using data from Lambda and estimates of the cost of count variability.

Chapter 8 examines practical strategies that hospital managers can use to reduce and mitigate inventory inaccuracy (Q3). Their managerial implications are discussed and a sequential and iterative improvement framework is proposed.

Finally, chapter 9 summarizes the conclusions and contributions of this dissertation, and provides suggestions for further research.

Chapter 2

The Hospital Supply Chain

Hospitals are a major and critical component of the health care delivery system in developed economies. In the U.S., hospital care (excluding services performed by physicians in a hospital but billed independently) accounted for 31% of total health care expenditures in 2007, representing the largest cost component (CMS, 2009).

This chapter presents background information on the hospital supply chain. First, the hospital supply chain is situated within the larger U.S. health care delivery system. Second, the hospital supply chain is viewed as socio-economic-technical system, and its stakeholders and technical infrastructure are presented.

2.1 The Hospital Supply Chain within the Health Care Delivery System

Throughout this thesis, descriptions about health care and hospital supply chains always pertain to the U.S. context, unless otherwise noted. This choice of focus stems from large differences in the organization and financing of health care delivery across countries, and in particular, the fact

that the U.S. does not have a universal health care system administered by government affiliated entities, unlike most developed economies (Institute of Medicine, 2004).

2.1.1 Overview of the U.S. health care delivery system

A health care delivery system is characterized by multiple types of stakeholders: Patients or the Public at-large, who require health care services; Providers, who deliver the care in different types of facilities (hospitals, clinics, pharmacies, nursing homes); and Payers, who reimburse Providers for the services performed, from funds collected as insurance premiums or taxes, as displayed in Figure 2-1. Charities are payers and/or providers. This describes, at a high level, the flow of care services and money.

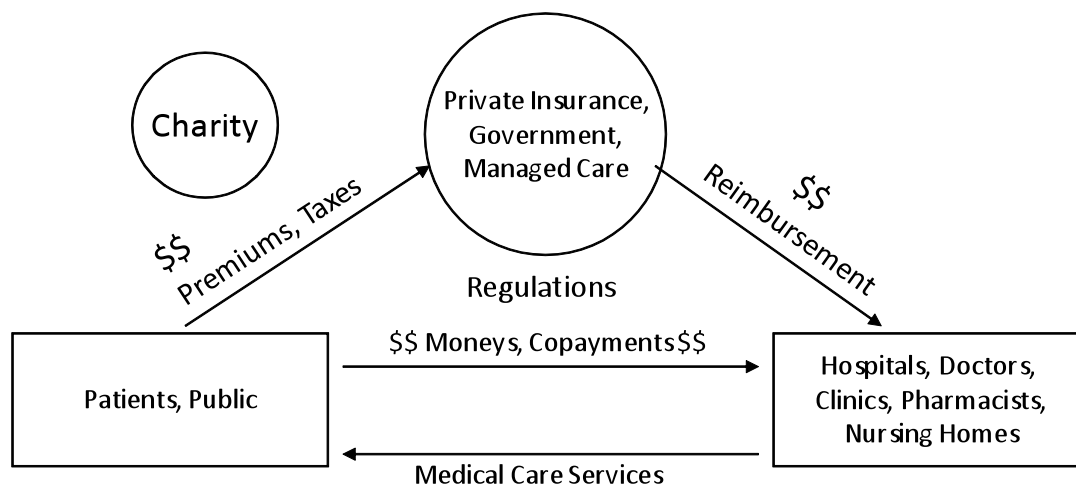


Figure 2-1: Flows of Money and Services in a Health Care Delivery system (Berndt, 2007)

Mentzer, et al. (2001) define a supply chain as “a set of three or more entities (organizations or individuals) directly involved in the upstream and downstream flows of products, services, finances, and/or information from a source to a customer.” The previous description of the health care delivery system is limited to the delivery of the care services, and should be completed with a description of the physical, financial and information flows related to the medical goods that enable care. Based on a three-year study of the health care value chain, Burns (2002) decom-

poses the chain into five steps defined by groups conducting similar tasks: Payers (government, employers, individuals), Financial intermediaries (insurers, health-maintenance organizations and pharmacy benefits managers), Providers (hospitals, doctors, outpatient clinics, pharmacists), Purchasers (distributors of pharmaceuticals and medical and surgical goods, and group purchasing organizations (GPOs)), and Producers (pharmaceuticals and medical devices manufacturers, medical surgical manufacturers, information technology vendors and capital equipment manufacturers).

2.1.2 The hospital within the health care supply chain

The flow of medical goods, information and money defines the health care supply chain, with products flowing from raw materials suppliers, through manufacturers and distributors, throughout the hospital until they reach the patient. From the perspective of the hospital, the health care supply chain can be decomposed into the external supply chain and the internal supply chain (Rivard-Royer, Landry, & Beaulieu, 2002). The external supply chain stops at the receiving dock the hospital. The internal supply chain starts at the dock of the hospital, comprises a storeroom and different storage areas on patient care units, and ends at the patient (Figure 2-2).

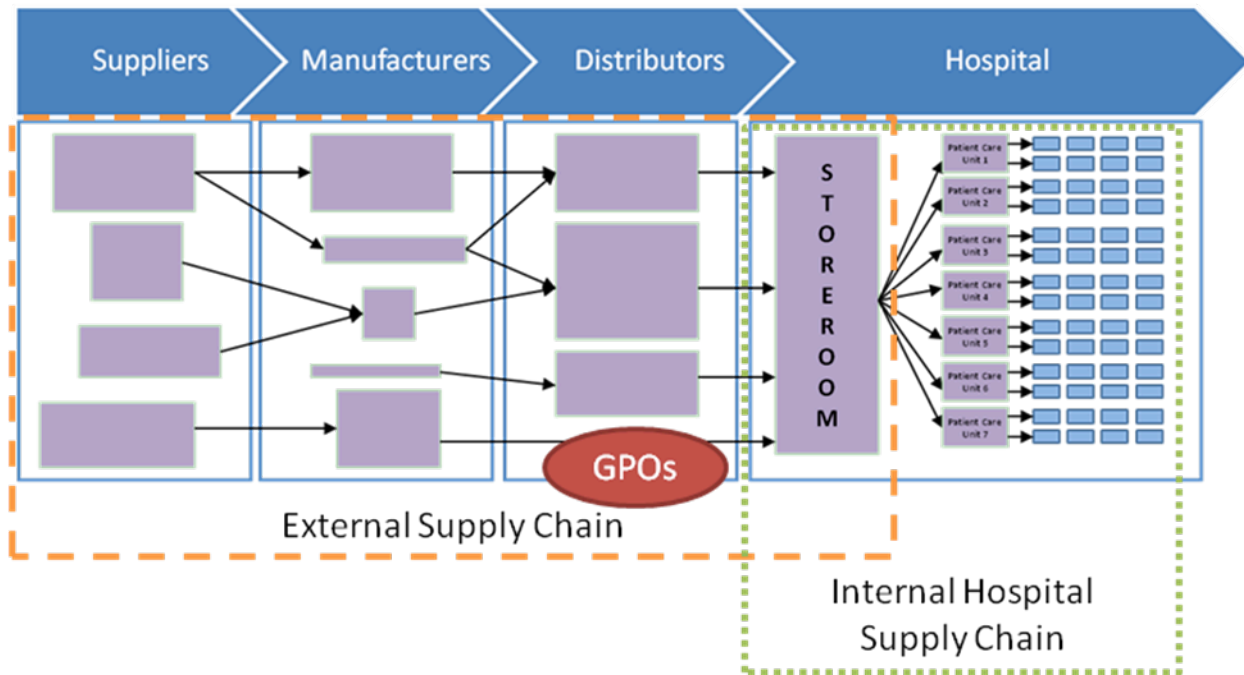


Figure 2-2: The Health Care Supply Chain (Oliveira, 2005)

While this dissertation focuses on the internal hospital supply chain, it is useful to have more information on the characteristics of external supply chain. At first sight, it follows a typical supply chain structure with products flowing from raw materials suppliers through manufacturers, distributors and finally to the hospital. The U.S. hospital industry comprised of 4,900 community hospitals¹ in 2007 (AHA & Avalere, 2008). With the notable exception of the Veterans Health Administration, which operates 153 medical centers and over 900 ambulatory care and community-based outpatient clinics (VA, 2009), the vast majority of hospitals are independently owned

¹ All nonfederal, short-term general and specialty hospitals whose facilities are available to the general public.

and operated². Such hospitals lack the scale to coordinate inbound logistics activities with the large numbers of medical goods manufacturers and consequently they rely on distributors to centralize such activities. Similarly, hospitals often join Group Purchasing Organizations to standardize and increase purchases of common products and achieve greater bargaining power in pricing negotiations (see Figure 2-2).

Conversely, drug distribution has become highly concentrated as a result of acquisitions (Finkelstein, 2003), with the top three players Cardinal Health, McKesson Corporation and AmerisourceBergen totaling a wholesale drug market share of approximately 90% in 2007 (Britt, 2007). The same players also dominate the distribution of commodity single-use medical supplies (e.g. needles, surgical pads). High-end and specialty single-use medical devices (e.g. cardiac stents) manufacturers often bypass distributors by relying on their direct sales forces to coordinate logistics activities, with product delivery being performed either by a sales representative when his presence is requested by the surgeon or by a third-party express carrier such as UPS or FedEx.

2.2 The Hospital Supply Chain as a Socio-Economic-Technical System

Bartolomei et al. (2006) define an Engineering System as “*a complex socio-technical system that is designed, developed, and actively managed by humans in order to deliver value to stakeholders*” (Bartolomei, et al., 2006, p. 3). This definition extends the boundary of the system

² In 2009, the Veterans Health Administration’s network was considerably larger than the largest for-profit hospital chains, such as Hospital Corporation of America (HCA) (165 hospitals, 112 outpatient centers), Tenet Healthcare Corporation (58 acute care hospitals) (Tenet, 2009).

beyond technical interactions traditionally considered by engineers to include social, political and economic interactions. This thesis views the hospital supply chain as an Engineering System under this definition, and seeks to describe and analyze it as such in order to inform and guide subsequent analyses.

In Decisions with Multiple Objectives, Keeney & Raiffa (1993) define objectives as indicators of the preferred direction of change; attributes as measurable quantities that indicate the extent to which an objective is achieved³; and goals as thresholds of achievement in a particular attribute (which therefore are either achieved or not achieved). In real-world problems, the set of objectives and attributes is not given *a priori*, and preferences for different objectives vary among different stakeholders (Keeney & Raiffa, 1993). The hospital supply chain is of course no exception to this statement. While it is the object of multiple objective decision analysis to quantify such preferences in order to recommend optimal decisions, this thesis seeks to determine the objectives and attributes of the system only as a lens through which to describe the hospital supply chain⁴.

³ According to these authors, attributes are *comprehensive* if “*knowing the level of an attribute in a particular situation the decision maker has a clear understanding of the extent that the associated objective is achieved*” (Keeney & Raiffa, 1993, p. 39). Attributes should be both *measurable* and *comprehensive* in order to be useful to the decision analyst.

⁴ As mentioned later in this section, objectives of health care systems often cited in the literature are improved quality of care, improved safety, and reduced cost (sometimes referred to as increased affordability or replaced with improved accessibility). Developing attributes of these objectives and assessing preferences among them constitute large research questions in health policy and health economics, beyond the scope of this thesis.

The next section describes the objectives of the hospital supply chain. The following section describes the different stakeholders of the hospital supply chain and highlights the objectives most relevant to each category of stakeholders.

2.2.1 Objectives

Keeney & Raiffa (1993) refer to MacCrimmon's (1969) suggestion of literature review, analytical study and "casual empiricism" as methods to elicit objectives. This thesis draws the objectives of the hospital supply chain (as an engineering system) from the literature, from empirical observations from semi-structured interviews with materials managers at different hospitals, and from field visits at Lambda.

The Institute of Medicine of the National Academies has proposed the following six aims for the health care system: Safety, Effectiveness (avoiding under use and overuse), Patient-centricity, Timeliness, Efficiency and Equity (America, 2001). Effectiveness is also sometimes referred to as Quality of Care and is a primary objective of the system. Efficiency is defined as avoiding waste of resources, time and ideas. Another objective is increased cost-effectiveness, defined as a ratio where the denominator is a gain in health according to a specific measure (e.g. years of life) and the numerator is the cost associated with the health gain (Gold, 1996, p. xviii).

According to annual surveys conducted by the American College of Healthcare Executives, financial challenges were cited within as a top concern by 77% of U.S. hospital CEOs, while patient safety and quality came in second, being cited by 43% of respondents (ACHE, 2009), as seen in Figure 2-3.

Issue	2008	2007	2006
Financial challenges	77%	70%	72%
Patient safety and quality ¹	43%	NA	NA
Care for the uninsured	41%	38%	37%
Physician-hospital relations	32%	35%	40%
Personnel shortages	30%	30%	30%
Governmental mandates	26%	22%	23%
Patient satisfaction	22%	17%	16%
Capacity	16%	11%	11%
Technology	9%	8%	8%
Issues about not-for-profit status	2%	4%	3%
Malpractice insurance	2%	2%	3%
Disaster preparedness ²	1%	1%	1%
Patient safety	NA	29%	27%
Quality	NA	33%	29%

Figure 2-3: Percentage of hospital CEOs citing issue among three most important (ACHE, 2009)

For the hospital supply chain, Quality of Care and Patient Safety objectives translate into an overarching objective of *delivering the drugs and medical devices and supplies⁵ as required for the care of the patient*. In the context of the hospital, doctors determine the required drugs and supplies necessary for patient care, and their cost is covered by different stakeholders depending on billing and reimbursement schemes. The issues surrounding choices of products by physicians for use on patients are of significant complexity and are the object of extensive study in the fields of health economics and health policy. Therefore, this thesis, while still considering the patient as a stakeholder in the supply chain, focuses on the physician as the decision-maker in matters of product choice within the context of hospital care.

This first overarching objective translates into several lower level objectives, which collectively reflect maximizing service delivery:

⁵ These terms are used in the sense defined in section 2.2.4, page 42.

- Maximize product availability at the point and time of clinical use. In the extreme, the objective is stated as “Never Run Out” of product (Martin, 2006).
- Minimize adverse medical consequences in the event that the product is not available at the point and time of clinical use. In such cases, nurses and doctors seek an alternate solution, such as:
 - o Searching for the same product in a nearby ward
 - o Requesting an emergency delivery of the product from the hospital’s central warehouse (a process called dispatching)
 - o Substituting with a functionally similar product.
- Maximize effectiveness at detecting and removing expired or recalled products

A second overarching objective is resource efficiency or minimizing waste: i.e. ensuring that the resources devoted to the hospital supply chain are minimized for a given level of service.

Finally, as for-profit corporations, the distributors and product manufacturers and certain hospitals have a profit objective and seek to provide a financial return to their shareholders.

2.2.2 Stakeholders

This section seeks to describe the different stakeholders whose interactions affect the hospital supply chain. The stakeholders of the hospital supply chain are presented in an order opposite to product flow: patient, nurses, doctors, materials management staff, purchasing management, group purchasing organizations (GPOs), distributors, supply chain automation vendors, product manufacturers, and regulatory agencies.

Patient: the patient is the ultimate beneficiary of the hospital supply chain, as the goods delivered by the hospital supply chain will be administered to her (in the case of drugs) or otherwise used during the episode of care. However, while in the hospital, the patient’s role is mostly passive as she is not the decision-maker in matters of product choice or delivery times. Ultimately, the pa-

tient is directly affected by and cares most deeply about the Safety and Quality of Care objectives.

Non-Physician Clinical Staff: Non-Physician Clinical Staff includes nurses, physicians' assistants, and surgical technologists⁶. Figure 2-4 shows current estimates for the national shortage of registered nurses, which is projected to continue in the future (AHA & Avalere, 2008; Biviano, Tise, & Dall, 2007).

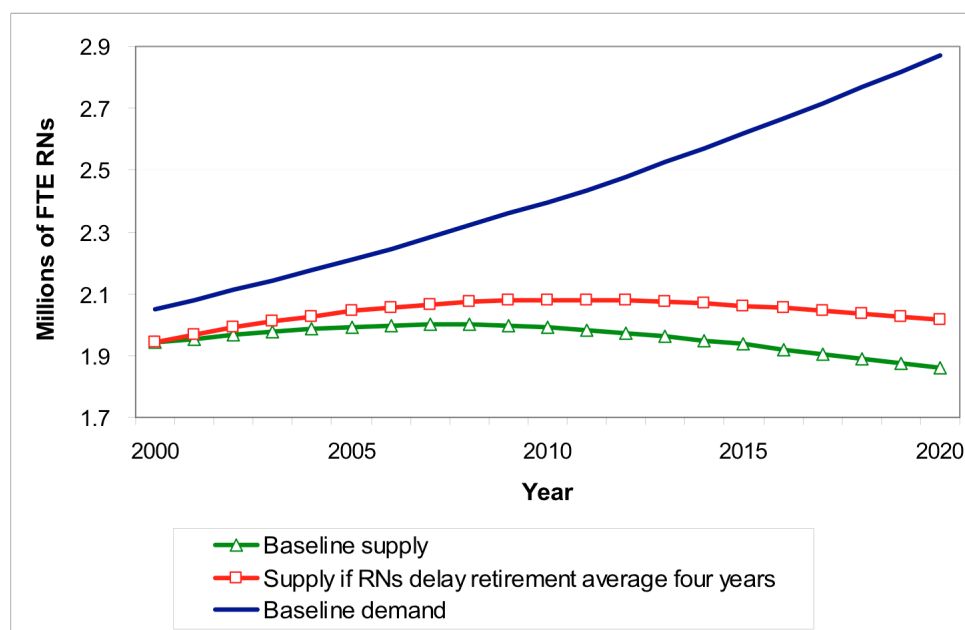


Figure 2-4: Projected full-time equivalent nurses demand and supply (Biviano, et al., 2007)

⁶ According to the Bureau of Labor Statistics' Occupational Outlook Handbook (2007), "Surgical technologists, also called scrubs and surgical or operating room technicians, assist in surgical operations under the supervision of surgeons, registered nurses, or other surgical personnel. Surgical technologists are members of operating room teams, which most commonly include surgeons, anesthesiologists, and circulating nurses."

This group is directly responsible for retrieving products from inventory as needed for patient care, based on physicians, surgeon or patient requests. Interviews conducted at Lambda, as well as a review of the literature, indicate that inventory management tasks are viewed as a time-consuming burden by non-physician clinical staff, standing in the way of their clinical mission (Hagland, 2005; Neil, 2003; Oliveira, 2005). In its marketing materials, a distributor of ADMs touts its automation product as “The Inventory Pain Reliever” and claims to “ease the pain and discomfort of inventory management” (Owens & Minor, 2005), playing on the negative attitudes held by clinical staff towards this task.

Physicians and Surgeons: Physicians and surgeons care most deeply about the availability of the product they prescribed or requested (as opposed to a substitute) at the time at which they need it. According to Sussman & Gupta (1992), “The physician’s primary concern is that the product functions effectively and appropriately at all times, at whatever cost. The purchasing manager’s concern is to make sure the product is available within acceptable, controllable costs”. In an operating room or a cardiac catheterization lab, the level of urgency implies that necessary supplies are prepared ahead of the procedure if possible, or promptly retrieved by a surgical technician or nurse.

This thesis defines end-users as all clinical staff including physicians, surgeons, nurses, physicians’ assistants, and surgical technologists.

Materials Management Technicians: Materials management technicians, sometimes called “refill technicians” are in charge of performing inventory management tasks in the hospital, such as receiving orders from suppliers, replenishing different stocking locations across the hospital, tracking expirations, and counting inventory periodically.

Purchasing Management: The purchasing department is in charge of procuring supplies and seeks to minimize costs. In order to gain leverage over manufacturers, many small hospitals join a Group Purchasing Organization, which conducts pricing negotiations on their behalf. Another

strategy commonly used is to develop a prime vendor agreement with a distributor. In a prime vendor agreement, the hospital commits to using the distributor for a period of several years to source pre-defined product categories. For the products covered by the agreement, hospital orders are fulfilled through the distributor, usually through Electronic Data Interchange. In exchange for its services, the distributor receives a service fee (usually a percentage of the product price). Prime vendor agreements seek to produce savings justifying the service fee through multiple mechanisms: (1) automation of labor-intensive tasks usually conducted by the purchasing department of the hospital; (2) lower inventory and transportation costs for the hospital; (3) lower unit prices obtained by reducing the number of suppliers used for a class of products (a process called standardization or value analysis), usually with a shift towards the distributor's brand (e.g. carry only one brand of hypodermic needles); (4) lower utilization of expensive products through improved reporting and on-site consulting services (Medline, 2009; E. S. Schneller & L. R. Smeltzer, 2006).

Group Purchasing Organizations (GPOs): Group Purchasing Organizations vary widely in their funding mechanisms, for-profit status, and the scope of the services provided to their members. Member fees (either flat rate or equal to a percentage of the dollar purchases of the member through the GPO) are a primary source of revenue for GPOs, as well as funds from vendors under a special anti-kickback statute exemption granted by the U.S. Congress in 1986⁷. Functions

⁷ There is a fierce policy debate about the overall benefit provided by GPOs to the healthcare system. Detractors of GPOs such as the Medical Devices Manufacturers Association oppose them on two main grounds: (a) member fees based on a percentage of dollar purchases are not an appropriate incentive for delivering cost savings; and (b) vendor fees violate the intent of the Medicare anti-kickback statute. A study funded by Health Industry Group Purchasing Association claims that GPOs provide \$36bn annually in price savings and \$2bn in avoided human resources costs (Schneller, 2009).

performed for members range from price negotiation and contract management to a broader array of services, including benchmarking and outsourcing of the hospital's purchasing department.

Distributors: Distributors play an essential role in linking over 20,000 suppliers to over 5,000 hospitals, and account for 80% of product flows to hospitals and health systems⁸ (E. S. Schneller & L. R. Smeltzer, 2006). Through their role as intermediaries and their network of warehouses, distributors offer both economies of scale and economies of scope in their logistics activities. Few product manufacturers are willing to take on the distribution function and manage the complexity of interacting with a large number of hospitals. From the perspective of hospitals, third-party distributors aggregate the supplies from many manufacturers in a single periodic shipment, simplifying ordering and inbound logistics activities.

Supply Chain Automation Vendors: These players provide supply chain automation solutions by selling, leasing and maintaining hardware and software products to support hospital inventory management, such as Automated Dispensing Machines (ADMs). The main players in this market are: CareFusion (CFN), a publicly traded subsidiary of major distributor Cardinal Health (CAH); Owens & Minor (OMI), a major distributor; and Omnicell (OMCL), a publicly traded company focused solely on this segment.

Product Manufacturers: Product manufacturers are focused on the production and marketing of medical supplies and devices. Their large network of sales representatives serves multiple purposes, including educating doctors about the benefits of their products, and in the case of surgical devices, providing training to clinical staff on product use. In some instances, especially for

⁸ Health systems, such as the Veterans' Administration or regional health networks, provide integrated care through a network of hospitals, outpatient clinics, and long-term care facilities.

high-value medical devices such as stents or implants, sales representative also have a distribution role, ensuring the delivery of the appropriate device for a particular medical procedure or monitoring consigned inventory.

Regulatory Agencies: There are three major regulatory bodies affecting the hospital supply chain: the Food and Drug Administration, the US Centers for Medicare and Medicaid Services (CMS), and hospital accreditation agencies. The Food and Drug Administration (FDA) is responsible for protecting the public health by assuring the safety, efficacy and security of drugs and certain medical devices⁹ (FDA, 2009), which it regulates mainly at the manufacturer and distributor level by mandating good manufacturing practices, and traceability in order to ensure that the supply chain has the capacity to execute product recalls. The Center for Medicare and Medicaid Services (CMS), part of the Department of Health and Human Services, oversees the Medicare and Medicaid public health care programs. Finally, to be eligible for Medicare reimbursement, hospitals need to be accredited by a hospital accreditation agency. While it is not the only accreditation agency, the Joint Commission has a partial monopoly on hospital accreditation in the United States, because in many states it is the only recognized accreditation agency valid for Medicare reimbursement eligibility. The Joint Commission¹⁰, formerly known as the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO), is a private not-for-profit organization dedicated to patient safety that inspects and accredits hospitals that has taken a leadership role in issuing safety recommendations and procedures for hospitals.

⁹ These terms are used in the sense defined in section 2.2.4, page 42.

¹⁰ Despite being a private organization, the Joint Commission has a partial monopoly on hospital accreditation in the United States, because in many states it is the only recognized accreditation agency valid for Medicare reimbursement eligibility.

2.2.3 Technical infrastructure

The technical infrastructure of a hospital's supply chain varies significantly across hospitals. This reflects the lack of standardization in the field and the reality that technology investment decisions made by hospital management occur dynamically over a number of years, with hospitals investing in ADMs typically to increase safety and/or reduce labor cost through automation.

The technical elements of a hospital's supply chain are both physical and information systems. Physical systems include fixed stocking locations such as shelves and cabinets as well as mobile components such as totes and carts. The purchasing information system tracks vendors, products, pricing information and orders. Common purchasing information systems vendors include Oracle PeopleSoft and SAP. The inventory information system tracks the available quantity of each product and reorder products and is complemented by physical components such as labels, barcodes, barcode scanners, inventory sheets, RFID chips and Automated Dispensing Machines. In contrast to other physical components common to any logistical setting, Automated Dispensing Machines (ADM) are specialty-purpose cabinets designed for storing drugs and medical supplies. Originally introduced to track the chain of custody around narcotic medications and avoid theft or abuse by hospital staff, point-of-use ADMs were used in 72% of U.S. hospitals in 2005 (Pedersen, et al., 2006). These electronic cabinets record the usage and inventory of medical supplies for inventory management purposes and can also be linked to billing and patient records.

Hospitals that do not use ADMs rely on daily inventory checks by materials management technicians, who identify items that are below a target inventory level (often this is done visually with some margin of error) and fill a pick list order using a paper form or electronic handheld device. Jayaraman, Burnett & Frank (2000) offer a detailed discussion of the operational issues prevalent in this context.

2.2.4 Product classification

The organization of the supply chain varies depending on the type of medical goods involved, and therefore it is necessary to distinguish between different types of medical goods. Different classifications schemes exist, including the North American Industry Classification System (NAICS), the Federal Supply Code used by the Military, and the risk-based Food and Drug Administration classification of medical devices, but there is no well-accepted, mutually exclusive and collectively exhaustive classification system applicable to all medical products. Definitions of what constitutes a “drug” or “medical device” vary across countries and regulatory agencies. The Federal Food, Drug and Cosmetic Act (Federal Food, 1938) defines a drug as:

“(A) articles recognized in the official United States Pharmacopoeia, official Homoeopathic Pharmacopoeia of the United States, or official National Formulary, or any supplement to any of them; and (B) articles intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease in man or other animals; and (C) articles (other than food) intended to affect the structure or any function of the body of man or other animals; and (D) articles intended for use as a component of any article specified in clause (A), (B), or (C). (...)” (Federal Food, 1938, [21 U.S.C. 321](g)(1))

The same act defines a “device” using a similar definition, to which is added the stipulation that it *“does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes”*. (Federal Food, 1938, [21 U.S.C. 321](h)(3))

This thesis classifies products into three categories: drugs (including biologics) (e.g. aspirin), single-use medical devices and supplies (e.g. an implantable stent, a catheter, or a set of sutures), and durable medical devices and equipment (e.g. an intravenous pump), recognizing the large differences in their supply chain characteristics. Drugs are subject to a specific set of FDA regulations and are the responsibility of the pharmacy once in the hospital. This segmentation is consistent with industry practice: Cardinal Health, a major distributor, divides its activities between

the “Pharmaceutical” segment, consisting of branded and generic drugs, and the “Medical” segment consisting largely of medical infection prevention/operating room products (Barrett, 2009).

Durable medical devices and equipment are managed very differently, as these constitute assets of the hospital as opposed to consumable products, and as such are excluded from further consideration in this thesis.

[Page intentionally left blank]

Chapter 3

Related Work

This chapter offers a brief overview of the academic literature on hospital inventory management, and then concentrates on the literature on inventory inaccuracy.

3.1 Hospital Inventory Management

3.1.1 A note on hospital supply chain management

Schneller & Smeltzer (2006) provide an in-depth survey of several hospitals selected for their commitment to strategically improving supply chain management. The book offers an explanation for the limited attention supply chains have received in healthcare until recently: they claim that materials managers have historically had limited power in organizations dominated by physicians.

A parallel can be drawn with the ascendance of Supply Chain Management to a C-level function during the 80's and 90's. Earlier the logistics function itself used to be fragmented (divided between inbound activities, outbound distribution, warehousing, inventory control and transportation management). Many companies today manage an integrated supply chain process, which includes these functions as well as procurement, supplier relationships, and in some companies,

manufacturing and customer service. In many hospitals, however, the concept of integrated supply chain management does not yet exist.

3.1.2 Hospital inventory management

Inventory management has historically received little academic attention compared to purchasing-centric themes in hospital supply chain management. An exception to this statement is the extensive work on blood bank inventory management, which is associated with the literature on perishable products.

Therefore, most of the literature is comprised of consulting reports and professional journals, such as *Materials Management in Healthcare* or *Hospital Material Management Quarterly*. These articles provide materials managers with information about operations management concepts and strategies (such as Six Sigma), IT and Automation solutions, and feedback from other hospitals on different approaches that have been applied. These publications are helpful in understanding the operational challenges faced by practitioners but are not targeted at an academic audience. Below is a brief review of relevant academic papers focusing on hospital inventory management, structured in chronological order.

Beier (1995) offers a descriptive, survey-based account of inventory management practices at hospital pharmacies in the U.S. in the mid-nineties. Because it focuses on pharmacies, it is important to note that it cannot account for the practices surrounding medical and surgical devices, or for the dramatically increasing use and cost of these devices in the past three decades.

Rivard-Royer, et al. (2002) discuss the evolution of those practices, and the advent in the late eighties of “stockless” schemes, where the hospital central storeroom was eliminated and delivery to patient care units was performed by the distributor. In some cases, such schemes were full Vendor-Managed Inventory implementations, where the distributor would determine orders on behalf of the hospital. Rivard-Royer and Landry report that “stockless” agreements eventually reached 10% of U.S. hospitals, before being abandoned by distributors who felt that the inventory savings at the hospital did not justify the increased replenishment costs, and that the fees

they received in exchange for replenishment services did not cover the costs of such programs. Different studies attempting to quantify the benefits and costs of such schemes are reviewed and the authors note that most of them are from potentially biased parties (e.g., distributors or consulting firms). The authors also offer a case study based on a Canadian hospital in Québec that adopted a hybrid approach, eliminating the storeroom for high-volume products by delivering case quantities to patient care units, and keeping a centralized inventory of low-volume items, which are broken into single units for delivery at the point of consumption. Through a combination of questionnaires and estimates of cost data, they compiled a breakdown of total costs before and after implementation. They conclude that the hybrid approach yields only marginal benefits to both the distributor and hospital, but that packing formats and storage areas can be redesigned to achieve substantial savings.

Recognizing that inventory management has been neglected in the healthcare sector and that normative methodologies are needed, Nicholson, et al. (2004) compared the inventory costs and service levels (as a proxy for quality of care) of two competing supply chain designs: (1) a three-echelon network comprised of the distributor, the central storeroom and patient care units versus (2) a two-echelon network typical of a “stockless” agreement that lacks a central storeroom. The analysis is focused on non-critical medical supplies and concludes based on numerical simulation of their model that the two-echelon network performs better both in terms of inventory costs and service levels. While the analysis is grounded in data collected through interviews, the focus on non-critical items and an additional assumption of low demand variability (coefficient of variation ranging from 0.3 to 0.4) restricts the applicability of the obtained results.

Lapierre, et al. (2007) derive a scheduling approach to optimize the logistics activities in a hospital, and explicitly aim to capture labor capacity constraints and utilization, in contrast with traditional inventory control models, which model distribution labor costs as fixed (per order) and/or variable (per unit) costs. Under the assumption of a constant demand rate during the planning horizon, a tabu search heuristic is provided to solve the formulated scheduling problem and numerical results are provided.

3.1.3 RFID applications

Radio-Frequency Identification (RFID) has been applied in hospitals for asset, patient and staff tracking applications as well as inventory management applications. While tracking applications have reported successful large-scale deployments, two factors limit the near-term applicability of RFID to reduce inventory inaccuracy in the hospital: (1) reliability, which is still problematic for items containing metals or liquids, and (2) the cost of adding an RFID tag on a consumable product, which is currently estimated at \$0.50 per box. For these reasons, such inventory management applications have so far been limited to high-value products, such as implantable devices.

3.2 Inventory Inaccuracy

This thesis defines inventory inaccuracy as a discrepancy between the book inventory (i.e., the inventory on hand recorded in the information system) and the physical inventory (i.e. the inventory on the shelf)¹¹.

This section first reviews empirical work assessing the magnitude and causes of inventory inaccuracies, then discusses models of the impact of inventory inaccuracy on service levels, and finally reviews models of optimized counting and ordering policies.

¹¹ Some authors use different terms, such as “recorded” for “book” and “actual” for “physical” (DeHoratius & Raman, 2008). This terminology is chosen to prevent confusion that could arise since this thesis is both empirical and theoretical: empirical analysis is performed on historical records (which some would call “actual”), while the discrete-event simulation model “records” (i.e. tracks) both book and physical quantities over time.

3.2.1 Empirical studies of Inventory Inaccuracy

Inventory inaccuracy has been documented in different sectors through empirical studies, ranging from government entities (Rinehart, 1960), military warehouses (Iglehart & Morey, 1972), utilities (Redman, 1995), manufacturing (Sheppard & Brown, 1993) and retail (DeHoratius & Raman, 2008; Gentry, 2005; Raman, 2000). The proportion of stock keeping units (SKUs) with inaccurate inventory records during counts varies widely across studies. These studies are difficult to compare directly due to different definitions of inventory inaccuracy (see (Schrady, 1970) for a review of the advantages and disadvantages of different definitions) but Table 3.1 provides an summary along with different drivers of inventory inaccuracy. One such driver is the frequency of inventory counts¹². To illustrate this point, consider the following: DeHoratius & Raman (2008) report that 65% of SKUs at a retailer had an inaccurate inventory record upon physical audit, while Klibanov & Eckel (2003) report that 19.5% of the drug Automated Dispensing Machines records were inaccurate. Both studies use the same definition of inventory inaccuracy, but the frequency of counts in the first study was annual or semi-annual, whereas the second study used much more frequent counts (approximately biweekly)¹³.

¹² Other terms used in the literature include “physical audits” and “inspections”.

¹³ As Klibanov & Eckel (2003) collected data for multiple products during a 10-day period and discrepancies were recorded and corrected on an ongoing basis. Therefore it is not possible to know the exact count frequency, but it is likely that counts are quite frequent.

<i>Study</i>	<i>Context</i>	<i>Definition of Inventory Inaccuracy</i>	<i>% of inaccurate SKUs</i>
Rinehart (1960)	Government agency	Absolute Discrepancy \geq 1% of inventory record balance; or Absolute Discrepancy * Unit Cost \geq \$1	Major items: negligible Secondary items: 20-50%
Emma (1966)	Military warehouse	Absolute Discrepancy > 24 units after one year	25%
Sheppard & Brown (1993)	Electronics Manufacturing	Absolute Discrepancy > 4 units after three months	27-36%
Ernst, Guerrero, & Roshwalb (1993)	(a) Retail pharmaceutical firm warehouse (b) Two finished-goods industrial manufacturing firm warehouses	Absolute Discrepancy > 0	(a) 83.7% (b) 91.9%, 95.5%
Millet (1994)	Logistics warehouse	Absolute Discrepancy > 0	36%
Klibanov & Eckel (2003)	Hospital, Automated Dispensing Machines for Drugs	Absolute Discrepancy > 0	8.9% - 26.4% Average = 19.5%
Kang & Gershwin (2003)	Retail	Absolute Discrepancy > 0 Absolute Discrepancy > 5 units	49% 24%
DeHoratius & Raman (2008)	Retail	Absolute Discrepancy > 0	65%

Table 3.1: Inventory Inaccuracy definitions and magnitude in empirical studies

Sources of inventory inaccuracies

The existing literature cannot be used to infer the relative magnitude of the sources of inventory inaccuracy across different industries due to the lack of standardization of counting policies.

Atali et al. (2006) define three sources of inventory inaccuracies: product misplacements, shrinkage and transaction errors. Product misplacements in the inventory record are also called physical errors (de Kok, van Donselaar, & van Woensel, 2008), because the book inventory is correct but the product cannot be found at its correct stocking locations (for instance, it is in a different aisle). Retailers define shrinkage as the sum of employee theft, shoplifting and vendor dishonesty, all of which result in depleted physical inventory without correspondingly updated book inventory. According to retail industry surveys, shrinkage accounted for 1.7% of sales in 2002 (Hollinger, 2003).

3.2.2 Effect of Inventory Inaccuracy on service levels

Morey (1985) provides an analytical model to examine the effect of increasing the frequency of cycle counts, increasing the safety stock or reducing inventory inaccuracies on the service level. Inventory inaccuracies are characterized only by their mean and variance per time period. Under the assumption of approximately normal lead time demand and independence of inventory errors with lead time demand, Morey (1985) derives the Minimum Average Protection Level (MAPL), an upper bound on the probability of stocking out during a replenishment cycle (i.e., the Type I service level). Analyzing how the MAPL evolves in response to changes in cycle count frequency, safety stock or inventory inaccuracy parameters allows managers to evaluate the tradeoff between these different management strategies. In a continuous review (R,Q) setting and under the assumption of normal lead time demand, Kumar & Arora (1992) quantify the effect of inventory inaccuracies on the on-shelf availability and find that lead time variability compounds their effect. Defining the relative inventory inaccuracy l as the ratio of the inventory error to the correct inventory, estimates of the in-stock probability are derived for different values of l , and numerical integration over the probability distribution function of l generalizes the model to stochastic inventory inaccuracies. Kumar & Arora (1992) apply the model to empirical data from a service parts organization and show that the realized service level is significantly below the service level implied by the initial safety stock calculation.

In order to assess the potential benefits of RFID technologies in a retail context where shrinkage is common¹⁴, Kang & Gershwin (2005) assume truncated normal demand and independent Poisson-distributed inventory errors and show, using a Monte Carlo simulation model, that even small rates of stock loss can create severe out-of-stock situations. In a multi-echelon context involving a producer, a distributor and a retailer, Fleisch & Tellkamp (2005) investigate the impact of different sources of inventory inaccuracy including theft, unsaleables¹⁵, misplaced items and incorrect deliveries for a high-value consumer packaged good through a Monte Carlo simulation. Using ANOVA, they report critical values of parameters characterizing sources of inventory inaccuracies at which performance measures differ significantly from the base case, but do not directly report the magnitude of the effect.

Table 3.2 summarizes the four models of the effect of inventory inaccuracies on service level described above.

¹⁴ Kang & Gershwin (2005) use the term “stock loss” for shrinkage.

¹⁵ i.e. “damaged, out-of-date, discontinued, promotional, or seasonal items that cannot be sold any longer”, which can be detected by the retailer or the end customer.

<i>Model</i> <i>Characteristic</i>	<i>Morey (1985)</i>	<i>Kumar & Arora (1992)</i>	<i>Kang & Gershwin (2005)</i>	<i>Fleisch & Tellkamp (2005)</i>
Inventory policy	Continuous review w/ safety stock (B)	Continuous review (R, Q)	Daily periodic review (R, Q)	Weekly review (R, Q)
Demand	i.i.d. Normal	i.i.d. Normal lead time demand	i.i.d. Truncated normal	i.i.d. Normal
Lead time	i.i.d. Normal and independent of Demand		Constant	Constant
Backlogged demand	No	No	No	Yes
Inventory inaccuracies	i.i.d. Known mean and variance ¹⁶	<ul style="list-style-type: none"> • Fixed relative inaccuracy • Probability distribution 	<ul style="list-style-type: none"> • Shrinkage: i.i.d. One-sided Poisson 	<ul style="list-style-type: none"> • Shrinkage: i.i.d. Uniform • Misplacement: i.i.d. Compound Bernoulli-Uniform
Counting policy	Periodic counts	None	<ul style="list-style-type: none"> • None • One count 	<ul style="list-style-type: none"> • Periodic counts • Correction on Stockouts¹⁷
Service level metric	Type I service level (MAPL)	Fill rate	Fill rate	In-Stock Probability
Methodology	Stochastic analytical model	<ul style="list-style-type: none"> • Stochastic analytical model • Numerical integration 	<ul style="list-style-type: none"> • Deterministic analytical model • Stochastic simulation 	Stochastic simulation

Table 3.2: Summary of the characteristics of models of the service level impact of inventory inaccuracy

¹⁶ Although no distributional assumptions are made, the normal approximation used by Morey (1985) implies that the distribution of errors is symmetric.

¹⁷ Fleisch & Tellkamp (2005) assume that book inventory is corrected during periodic counts and whenever the physical inventory reaches zero, effectively assuming an instantaneous Zero-Balance Walk (ZBW) policy.

3.2.3 Counting and ordering policies under Inventory Inaccuracy

A stream of literature aims to characterize improved or optimal counting and/or ordering policies in the context of inventory inaccuracies. The first part of this section focuses on counting policies, then the section presents work on ordering policies.

Counting policies

Inventory counts serve two purposes: (1) recording assets for accounting and tax reasons; and (2) reducing inventory inaccuracy. Counts performed for accounting and tax reasons are usually conducted as physical inventories: the entire facility is shutdown and all products and locations are counted at once, at yearly or quarterly intervals. For this reason they are less frequent than cycle counts, whose purpose is to improve inventory accuracy by counting only a sample of products at a time.

The literature uses the term “cycle counting” to cover a wide array of counting methodologies, and a brief review of the literature on these methods is provided in Rossetti, et al. (2001), with a more in-depth treatment in Brooks, et al. (2007) and Piasecki (2003). This thesis restricts the definition of *cycle counting* to sampling methods that are independent of the inventory state. This corresponds to the traditional definition, where items are assigned a target count frequency depending on product characteristics (Kok & Shang, 2009). On the other hand, counting policies that depend on the state of the book or physical inventory are thereafter referred to in this thesis as *dynamic counting policies*.

Cycle counts are increasingly being used to address inventory accuracy challenges, with the proportion of retailers intending to perform “more cycle counts” rising from 46% to 78% between 2001 and 2003 (CSA, 2001, 2003). In fact, the proportion of retailers who performed more than ten cycle counts per year increased from 25% in 2004 to 37% in 2005 (Foundation & BearingPoint, 2005).

Several papers focus on determining the optimal count frequency. One study (Iglehart & Morey, 1972) considers the case of a single-item, periodic review inventory system with a pre-

established (s,S) ordering policy. The authors assume that the only source of inventory inaccuracy is transaction errors, and do not consider misplacement or theft. They jointly derive the optimal count frequency and additional buffer stock required to minimize total cost per unit time, subject to a service level constraint, both in the case of perfect and imperfect audits. Morey & Dittman (1986) derive the minimum frequency of physical audits that keeps inventory record inaccuracy (rather than the service level) within a set limit for a single item.

Dynamic Ordering and Counting policies

Lee & Özer (2007) examine the literature on Radio Frequency Identification (RFID). They highlight the importance of deriving ordering policies that account for inventory inaccuracy in order to have a correct benchmark against which to value technologies claiming to eliminate inventory inaccuracies. Using a dynamic programming formulation, Atali, et al. (2006) derive an optimal base stock policy, with the base stock level depending on the time elapsed since the last audit. Kok & Shang (2007) show that for inventory inaccuracies that have mean of zero, an inspection adjusted base stock policy is optimal for the single-period problem, and propose a cycle-count adjusted base stock policy that modifies the base stock level according to the time since the last cycle count. DeHoratius, et al. (2008) introduce the notion of a probabilistic Bayesian inventory record and show how it can be stored, updated, and then used to derive dynamic ordering and counting policies.

[Page intentionally left blank]

Chapter 4

Problem Statement and Approach

This chapter first states the problem and defines the research questions. It then presents the research approach and study site.

4.1 Problem Statement

This study takes a holistic perspective on the hospital supply chain and recognizes the social, technical and managerial components of this Engineering System in order to improve product availability in the hospital (Figure 4-1).

The Social component acknowledges that the system is comprised of multiple stakeholders whose behavior is affected by their own values and objectives. This system is also Technical, using information technology and a large physical infrastructure of stocking locations to coordinate the procurement, receiving, delivery and dispensing of products. Lastly, a Managerial component drives the system behavior by setting policies, providing ongoing feedback and training.

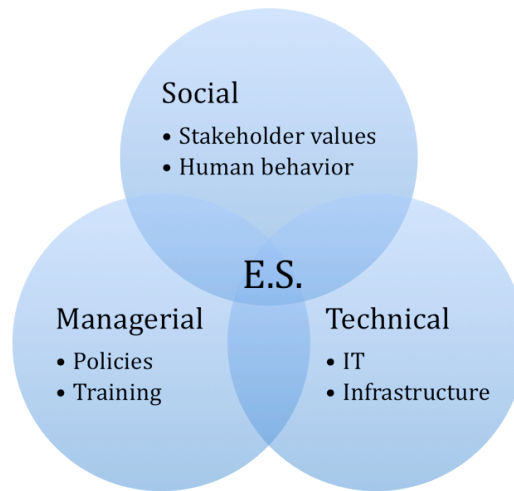


Figure 4-1: Conceptual view of the Social, Technical and Managerial components of the Hospital Supply Chain as an Engineering System

4.1.1 Discussion of the literature

The academic literature on hospital inventory management is heavily focused on comparing the cost performance of different supply chain designs (e.g., pooled inventories in a central warehouse vs. “stockless” schemes with direct deliveries to the wards). Unlike retail, stock out costs at hospitals are essentially intangible and unrelated to the product value, but manifest themselves in the form of additional nursing labor, managerial effort and patient risk (Tucker, 2004). For this reason, the focus of this thesis is on product availability. However, it is important to ensure that proposed solutions are practically feasible, and avoid raising operational costs.

While Klibanov & Eckel (2003) report that 19.5% of a hospital’s medication inventory records were inaccurate, the type and number of inventory counts that were conducted during the study

period are not provided and therefore it is difficult to put this figure in context¹⁸. More generally, knowledge about the sources and patterns of inventory inaccuracies in a hospital setting is limited.

The literature has examined the effect of inventory inaccuracies on product availability through analytical and simulation methods, however the following factors have not been investigated:

1. The effect of inventory inaccuracies on slow-moving items for which the normal demand distribution assumption (and more generally a continuous distribution) is not valid;
2. The impact of inventory inaccuracy on *book* (i.e., measured) vs. *physical* (i.e., actual) service level metrics;
3. The role of counting policies in mitigating high levels of shrinkage (i.e. > 5% of demand).

4.1.2 Research questions

The limitations of the research previously described suggest the following research questions:

(Q1a) *Can we identify and quantify specific user behaviors that introduce inventory inaccuracies in the hospital?*

(Q1b) *How are inventory inaccuracies currently mitigated in the hospital?*

¹⁸ Klibanov & Eckel (2003) focus on medication errors and therefore inventory management is not of direct interest to their study. The paper does not suggest that all items were inspected for discrepancies.

(Q2a) *What is the quantitative impact of inventory record inaccuracies on book and physical product availability?*

(Q2b) *How do counting policies affect product availability?*

(Q3) *What are effective strategies that hospital managers can use, recognizing the interdependencies between different strategies?*

4.2 Research Approach

The research is empirically grounded by a longitudinal case study of a New-England area hospital, Lambda, which is presented in subsection 4.2.1. Research questions (Q1a) and (Q1b) are approached using interviews and data from Automated Dispensing Machines described in subsection 4.2.2. Research questions (Q2a) and (Q2b) are explored using a Discrete-Event Simulation model and an analytical model. Finally, research question (Q3) is addressed through an integrated analysis including the exercise of the DES and analytical models on real-world data.

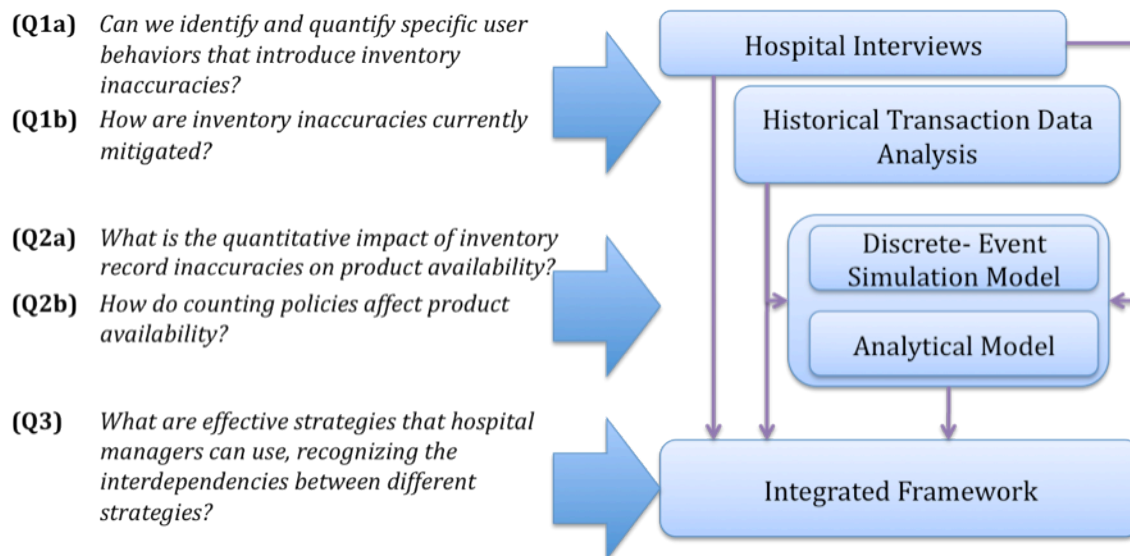


Figure 4-2: Diagram summarizing the multi-method approach to the research questions

4.2.1 Research site and scope

The research site is a 300-bed New England hospital, referred to here as Lambda, with annual inpatient admissions of approximately 20,000 patients, revenues of approximately \$800m, an operating margin of 3% and a total staff of 5,000 employees. Purchased supplies and services as a percentage of costs is in line with the industry average at 31% (E. S. Schneller & L. R. Smeltzer, 2006).

In addition to the quantitative data described in the next section, semi-structured interviews were conducted at Lambda and other hospitals to gain a detailed understanding of inventory management processes (the list of the titles of persons interviewed is provided in Appendix A).

4.2.2 Automated Dispensing Machine field data

Lambda has deployed a network of 108 Automated Dispensing Machines across its 35 wards for medical and surgical supplies. Lambda also uses Automated Dispensing Machines for medica-

tions dispensing, but drugs are tightly managed (at Lambda as well as other hospitals, for regulatory reasons) through the pharmacy, and were not considered in this study. Each machine, or station, comprises one or multiple closed-door cabinets and drawers storing medical and surgical supplies. To retrieve supplies, perform replenishment or inventory counting, a nurse or replenishment technician can log in into a console, using a password or biometric identification. Additionally, some products are available on open shelves, and their use can be recorded by pressing a radio-frequency button that is linked to the station.

Transaction-level demand, replenishment, inventory policy, inventory audit data at the Station-SKU level, covering a period of 15 months, were extracted from the Automated Dispensing Machines (ADMs) database and stored in a MySQL database. This dataset consists of 3.6 million transaction records containing multiple fields: a unique identifier for the Station, the Stock Keeping Unit (SKU) number, identifiers of the drawer in the cabinet where the SKU is stored, the type of transaction (Demand, Return, Replenishment or Count), the quantity involved in the transaction, the expected inventory prior to the transaction, the physical inventory at the beginning of the transaction, the inventory at the end of the transaction, the reorder point, and the order-up-to level (see Table 4.1 for a sample transaction record).

Station	S_CL_A
Transaction Time	2/1/2007 16:19
Transaction Type	Demand
Drawer	56
Pocket	26
Supplyclass	4
Item ID	61367
Item Name	KWU2014W23
Quantity	1
Expected Count	2
Actual Count	2
End Count	1
Minimum (Reorder point)	3
Maximum (Order up to level)	5

Table 4.1: Sample Transaction Record for Item 61367 in Station Cath Lab A

[Page intentionally left blank]

Chapter 5

Characterizing and Quantifying Sources of Inventory Inaccuracies

This section provides a typology of the different sources of record inaccuracies in Automated Dispensing Machines. Some of these sources are analogous to the causes of record inaccuracies in a retail environment, while others are specific to the hospital context. Sources of record inaccuracies are usually highly socio-technical, as they are often driven by both behavior and technological factors. The relative magnitudes of the different sources are evaluated in the hospital environment. Inventory record inaccuracies are a subset of record inaccuracies, where record inaccuracy is defined as the discrepancy between the records in an information system and the true state of the system. Record inaccuracies can also affect quantities other than inventory, such as sales, product returns, and replenishment quantities.

After classifying the potential sources of inventory inaccuracies, the most common ones are analyzed qualitatively and quantitatively using Lambda's data.

5.1 Typology of Sources of Inventory Inaccuracies

In a retail context, DeHoratius & Raman (2008) mention different sources of inventory inaccuracies based on the inventory management processes during which they occur: Selling and Re-

stocking, Replenishment, Database Synchronization, and Inventory Counts. They also identify product theft as a potential source of inventory record inaccuracy.

This section distinguishes between the following sources of inventory inaccuracies at Lambda: Imperfect demand recording, incorrect returns, incorrect replenishments, incorrect counts, and fake transactions.

5.1.1 Imperfect Demand Recording

Imperfect Demand Recording occurs when the end-user withdraws products from a machine and product usage is either not recorded at all or is recorded inaccurately. Multiple causes of Imperfect Demand Recording exist, including behavioral and technical issues.

- a) The end-user forgets to press the “Take” button to record the transaction (behavioral issue).
- b) The end-user withdraws multiple units of a product but only presses the “Take” button once, thereby not keying in the correct quantity withdrawn (behavioral issue).
- c) The end-user presses the “Take” button, but the button fails to record the transaction because of a mechanical malfunction or a dead battery. Some malfunctions can be intermittent, making them difficult to detect (technical issue).

From an inventory record perspective, such inaccuracies are analogous to product theft in a retail setting. Without the existence of a checkout counter, products are withdrawn without the inventory record being updated.

During interviews, inventory managers at Lambda defined “Compliance” as the act of properly recording the quantity withdrawn (i.e., (a) and (b) above are considered non-compliant), and they considered non-compliance as the main source of inventory inaccuracies. Improvement efforts have consisted of educating end-users about the necessity of pressing the “Take” button as they

withdraw product from the machine. For instance, stickers are posted on the Automated Dispensing Machines to remind end-users to press the “Take” button (Figure 5-1).



Figure 5-1 A “For patients’ sake, Press Take” sticker affixed to an Automated Dispensing Machine at Lambda hospital

5.1.2 Incorrect returns

Product returns occur when the incorrect product is withdrawn and then returned, or a product is not opened and therefore is not used on a patient. For instance, during a cardiac catheterization procedure, the physician may want multiple catheters of different sizes on hand, but will typi-

cally not use all of them. After the procedure, the catheters are returned to the Automated Dispensing Machine, and a “Return” button next to the “Take” button should be pressed. However, the end-user sometimes forgets to press the “Return” button or forgets to input the correct product quantity when returning more than one unit.

This can result in a situation in which the book inventory record understates the on-hand physical inventory. Recorded product returns represented 4.1% of all transactions, 5.3% of demand transactions, and 4.8% of demand quantities (see Table 5.1).

5.1.3 Incorrect replenishments

Replenishment transactions take place through the main console of the Automated Dispensing Machine, and are performed by a materials management technician. Replenishment problems occur when such a transaction is not recorded or the recording itself is inaccurate.

During interviews, the following scenario was described. A technician arrives with a product with which to refill the machine, but the console is being used by an end-user, such as a nurse searching for a product. Rather than wait for the end-user to clear the station as instructed, the technician may instead leave the product on the shelf without recording the replenishment through the ADM console. This issue can be potentially exacerbated by pressure from end-users, who may be waiting for the technician to clear the machine or the area altogether. Another possible scenario is one in which the replenishment technician finds too much inventory at the designated location for the product, and therefore lacks the space to “put away” the product. He may choose to leave the product on the floor or on an empty shelf without recording it. Such variations to the designated process can result in the book inventory understating the physical inventory and cause loss of product or trigger unnecessary replenishments.

Another source of errors during replenishment is keying errors. When a transaction quantity is input using a numeric keyboard, typing errors can occur and result in incorrect quantities. For in-

stance, “22” may be recorded instead of “2” or vice versa, and “6” may be entered incorrectly instead of “9”.

5.1.4 Incorrect counts

During a transaction, each user has the opportunity to correct the inventory record. This is the principal purpose of Count transactions. A user inputting an erroneous quantity for the inventory present on the shelf may fail to correct or introduce inventory inaccuracy.

5.1.5 Fake transactions

End-users sometimes try to modify the system’s behavior in unorthodox ways. For instance, the following pattern was discussed during the interviews. In a ward experiencing a high number of stock-outs during weekends as a result of reorder points that were too low, a particular nurse recorded artificial discrepancies or fake demand transactions on Thursday to trigger a large delivery on Friday morning, to cover for the weekend demand. However, most of these cases took place in the early phases of implementation and no current evidence (qualitative or quantitative) exists that such behavior is still frequent.

Another source of fake transactions came from the necessity of having a metric for imperfect demand recording. The ADM system records a “Null transaction” whenever an end-user logs into the ADM, opens a door cabinet, and shuts the door without pressing the “Take” or “Return” button. In theory, this can happen because either:

- a) the end-user forgot to record the product withdrawal or return, or
- b) no product was withdrawn or returned (for instance, the product sought by the end-user was not found behind this particular cabinet door).

In order to use Null transactions as a proxy for Imperfect Demand Recording, end-users at Lambda were instructed to press the Take button followed by the Return button if they were in

scenario (b) above. This ensures that Null transactions only correspond to scenario (a)¹⁹. However, Null transactions cannot be tied to a specific product (since a given cabinet door often contains multiple products) and therefore were not directly usable at the Station-SKU level.

5.2 Existing Metrics of Inventory Inaccuracy

Automated Dispensing Machines (ADMs) calculate and report different Compliance metrics from the recorded transaction data. These metrics are aggregated at the Station level (i.e., they reflect all products stocked in a given machine) and are reviewed by materials management staff, usually monthly. First, the content of the “Null transactions” and “Compliance” reports is described, and then the limitations of these reporting mechanisms and metrics are analyzed.

5.2.1 Inaccuracy detection

During each transaction the user has the opportunity to correct inventory record inaccuracies, but in practice this does not occur, with over 95% of the discrepancies corrected during a Replenishment or Count transaction (Table 5.1). Interviews confirmed that nurses do not view correcting discrepancies as part of their responsibility and they lack the time to compare the physical inventory present on the shelf to the book inventory.

¹⁹ In theory, this could artificially inflate the number of Demand and Return transactions, but the software automatically discards transactions in which the Return button is pressed soon after the Take button. This was verified by filtering real-world transaction data. All unit-sized Return transactions that followed a unit-sized Demand transaction within a one-minute interval were counted and amounted to 2,275 transactions out of a total of 147,195 Return transactions (1.55%).

Transaction Type	Demand	Replenish	Return	Count
Number of transactions	2,768,226	369,934	147,195	312,589
<i>% of transactions</i>	<i>76.94%</i>	<i>10.28%</i>	<i>4.09%</i>	<i>8.69%</i>
Total transaction quantities	12,032,720	12,875,764	577,319	11,032,984
Number of discrepancies	2,902	80,816	671	65,133
<i>% of discrepancies</i>	<i>1.94%</i>	<i>54.05%</i>	<i>0.45%</i>	<i>43.56%</i>
<i>Discrepancy detection rate (discrepancies/transactions)</i>	<i>0.10%</i>	<i>21.85%</i>	<i>0.46%</i>	<i>20.84%</i>

Table 5.1: Discrepancy detection rates for different transaction types

5.2.2 Existing reports

The Null transactions report counts the number of Null transactions by End-User (as identified by his or her biometric login) and by Station, ranked in decreasing order (Figure 5-2). This allows materials management technicians to focus their training efforts on individuals who have a high number of null transactions. In particular, newly arrived staff members who did not receive training on how to use the ADM system are often identified in this way. However, once “outlier” stations or end-users have been identified and addressed, this type of report offered limited additional insights that could help reduce Imperfect Demand Recording.

The Compliance report provides for each Station on a monthly basis:

- Number of discrepancies: number of times a mismatch between the book and physical inventory was corrected, across all transactions;
- Discrepancy quantities: the sum of the absolute value of the difference between the book and physical inventory records when discrepancies were detected, across all transactions;

- Number of transactions: the total number of transactions that were Demand or Return transactions;
- Transaction quantities: the absolute sum of the Demand or Return transaction quantities;
- Activity Compliance =
$$\text{Number of Transactions} / (\text{Number of Transactions} + \text{Number of Discrepancies});$$
- Quantity Compliance =
$$\text{Transaction Quantities} / (\text{Transaction Quantities} + \text{Discrepancy Quantities});$$
- Number of Stock-outs;
- Number of Null Transactions.

5.2.3 Limitations of Existing Metrics

Activity Compliance and Quantity Compliance attempt to characterize the amount of inventory inaccuracy in the system. Both of these metrics suffer from several limitations.

Aggregation bias

First, observations made at the Station-SKU level are implicitly weighted according to the number of demand / return transactions (for Activity Compliance) or the total demand / return quantities (for Quantity Compliance). This is of concern because fast-moving products may exhibit different compliance characteristics and are not necessarily the most critical ones.

Sensitivity to the number of counts

Second, these metrics do not adjust for the number of inventory counts performed at the Station during the month. This means that one can measure higher Compliance because fewer counts were performed and fewer discrepancies were discovered, whereas the accuracy actually decreased (described in Figure 5-4 as “Good QC without verified counts”). Count transaction data show that count frequency varies substantially across stations, products and time, therefore making comparisons and tracking these quantities over time is difficult at best and likely speculative.

Station	Num of Discrepancies	Discrepancy Quantities	Num of Transactions	Transaction Quantities	Activity Compliance	Quantity Compliance	Num of Stockouts	Num of Null Transactions
S_5W	543	11,217	8,320	46,187	93.87%	80.46%	58	241
S_6C	645	10,991	11,632	50,697	94.75%	82.18%	49	233
S_6E	366	6,526	7,261	32,618	95.20%	83.33%	48	720
S_6E_SYR	0	0	0	0	100.00%	100.00%	0	1
S_6SE	366	6,455	5,884	29,373	94.14%	81.98%	74	225
S_6W	561	9,895	6,801	33,454	92.38%	77.17%	53	474
S_7C	581	112,793	10,215	39,169	94.62%	25.78%	105	275
S_7E	509	12,063	10,961	29,862	95.56%	71.23%	120	656
S_7SE	388	8,088	5,404	24,650	93.30%	75.29%	71	306
S_7W	548	55,493	8,040	37,795	93.62%	40.51%	72	365
S_ANES	325	5,825	380	4,804	53.90%	45.20%	17	123
S_ANGHO_1	170	1,020	2,113	2,994	92.55%	74.59%	24	204
S_ANGHO_2	168	924	2,174	3,182	92.83%	77.50%	28	195
S_CCU_1	48	1,174	1,880	4,568	97.51%	79.55%	28	163
S_CCU_2	249	8,005	6,330	38,880	96.22%	82.93%	74	721
S_CL_A	502	2,606	3,188	4,865	86.40%	65.12%	131	197
S_CL_B	421	2,090	2,226	3,408	84.10%	61.99%	82	267
S_CL_C	246	2,203	1,625	2,384	86.85%	51.97%	54	161
S_CL_HOLD	23	893	106	2,506	82.17%	73.73%	11	8
S_CL_SR_1	201	1,454	1,735	13,257	89.62%	90.12%	190	436
S_CL_SR_2	118	2,851	797	6,157	87.10%	68.35%	49	163
S_ED_ENT	4	33	16	117	80.00%	78.00%	1	65
S_ED_IV	83	824	2,440	6,220	96.71%	88.30%	8	167
S_ED_MINOR	154	2,450	1,369	7,192	89.89%	74.59%	8	192
S_ED_NS_1	28	1,549	697	3,098	96.14%	66.67%	6	58
S_ED_NS_2	3	30	38	44	92.68%	59.46%	0	18
S_ED_RML_4	6	20	10	28	62.50%	58.33%	1	8
S_ED_RML_5	3	30	53	79	94.64%	72.48%	0	33
S_ED_SUT	4	27	102	182	96.23%	87.08%	0	35
S_ED_TELE	55	737	646	2,944	92.15%	79.98%	0	65
S_ED_TRA_1	41	1,444	117	271	74.05%	15.80%	9	105
S_ED_TRA_2	22	187	81	300	78.64%	61.60%	7	83
S_ED_UTIL	355	10,218	4,170	46,097	92.15%	81.86%	51	204

Figure 5-3: Sample Compliance Report

Sensitivity to outlier observations

Furthermore, Quantity Compliance received the most attention because it was recommended by the ADM vendor as the metric to focus on. However, this metric was sometimes abnormally low (< 50%), and materials managers discovered that it was not robust to outliers. For instance, a single keying error affecting only one product (recording a demand quantity as 666,666 instead of 6) could be immediately corrected by the user (creating a discrepancy of 666,660) and artificially lower the value of the metric (described in Figure 5-4 as “Bad QC with keying errors”). Materials managers often screen the transaction log for these erroneous transactions, make a note of them and manually (with a paper and calculator) recalculate the Quantity Compliance metric.

Quantity Compliance Goal > 95%

- 1.) Good QC ***with*** verified counts
- 2.) Good QC ***without*** verified counts
- 3.) Bad QC ***with*** refill keying errors
- 4.) Bad QC ***without*** refill keying errors

The goal is to reduce/eliminate # 2 & # 3.

Figure 5-4: Internal slide used at Lambda during materials management technicians training, illustrating the sensibility of Quantity Compliance (QC) to the number of counts and outlier observations

5.3 The Role of Framing

One material manager at Lambda mentioned that he reported the Quantity Compliance metric, rather than the Activity Compliance metric, because it had lower values (e.g., 81% instead of 94%), and made inventory inaccuracy look like a more important problem when reported to end-users. This anecdote suggests that the scale used for the reporting metric is important, and is consistent with the literature on the effect of the framing of outcomes on decision-making. In a seminal *Science* paper, Tversky & Kahneman (1981) presented subjects with a choice between two treatment programs for an upcoming disease that will save 200 lives on average, the first one with certainty and the second one by having a one-third chance of saving 600 lives and a two-thirds chance of saving no one. Mathematically identical choices were framed as either lives saved or projected deaths. When the problem was framed in terms of lives saved, 72% of subjects chose the first (risk-free) program, whereas in the “projected deaths” framing, 78% of subjects chose the second (risk-taking) program.

Therefore, it is plausible that the way a given metric is calculated and reported affects the organization’s response to it. For instance, Quantity Compliance is expressed on a percentage scale, with the ideal value at 100%. At the time interviews that were conducted, its common range at Lambda was 70–95% and no explicit target was set beyond the ideal value of 100%. A Quantity Compliance of 95% was considered a good outcome. This framing can be contrasted with the Defects Per Million (DPM) measure of process performance, where the scale is inverted (i.e., a lower number is better, with zero being the ideal value) and widened (i.e., small differences are large numbers). While a 95% Compliance is mathematically equivalent to 50,000 DPM, the second framing sounds much less favorable.

One could argue that managers choose metrics that present current performance in a more favorable light. At Lambda, this did not appear to be the case, as materials managers were facing the situation and trying to reduce Imperfect Demand Recording by end-users. They used and communicated the Compliance metrics reported built into the ADM software. This suggests that sig-

nificant improvements are possible if these behavioral effects are recognized and integrated into the design of information and reporting systems.

5.4 Quantifying Imperfect Demand Recording

The goal of this section is to create a metric of Imperfect Demand Recording that can be used for measurement and modeling. When estimated from ADM data and aggregated over different products, the metric can be tracked over time and can provide feedback on process improvement efforts to managers and end-users. Imperfect Demand Reporting should also be usable as an input parameter in a quantitative model.

This thesis defines Recording Accuracy as the percentage of actual product usage (i.e., sales net of product returns) recorded. For instance, if three units of products are withdrawn from an ADM cabinet and one unit of product is recorded as used, the Recording Accuracy is 33.3%²⁰.

$$\text{Recording Accuracy} = \frac{\text{Recorded Usage}}{\text{Actual Usage}}$$

$$\text{Discrepancy} = \text{Recorded Inventory} - \text{Actual Inventory}$$

$$\text{Recording Accuracy} = \frac{\text{Recorded Usage}}{\text{Recorded Usage} + \text{Discrepancy}}$$

²⁰ As defined, a Recording Accuracy above 100% is possible if the recorded quantity is overstated. Although this can happen, it is relatively rare in a hospital setting.

5.4.1 Challenges to estimating Recording Accuracy from ADM data

To estimate Recording Accuracy, one needs to compare the book and physical inventory records over a given measurement period. However, the ADM transaction record only contains the book inventory record and any discrepancies that have been noted between it and the physical inventory record. Discrepancies are detected during Count and Replenishment transactions (see Table 5.1). The absence of a discrepancy between the physical inventory record and the book inventory record is only meaningful if it is certain that a count was performed. This can be assumed in the case of Count transactions but not for Replenishment transactions.

Estimating Imperfect Demand Recording from current ADM data is challenging for the following three reasons.

a) Irregular count frequencies:

Count events are unevenly spaced across time and the time interval between two counts typically spans multiple estimation periods (e.g., monthly measurements).

b) Small sample sizes:

Counts for a single product may be infrequent relative to the estimation period. As a result, and all other things being equal, increasing the estimation frequency will result in smaller sample sizes and, therefore, wider confidence intervals around the estimates.

Additionally, low-demand items exhibit both low demand and discrepancy figures (usually less than five items used between two counts), resulting in wider confidence intervals for the estimate of the Recording Accuracy.

c) Heterogeneity:

Imperfect Demand Recording varies over time and across different products.

For instance, items accessible on an open shelf as opposed to a locked cabinet had a lower Recording Accuracy.

Pooling data for different products within a single period and accounting for heterogeneity can help mitigate these issues. Section 5.4.2 introduces the notations and estimation methodology for a single product. Section 5.4.3 considers the case of multiple products and allows for heterogeneity across different products. Finally, section 5.5 presents empirical results related to Recording Accuracy at Lambda.

5.4.2 Estimation methodology for a single product

Throughout this section, the following notations are used:

- T represents the number of periods considered;
- $t \in [1, T]$ represents the index for the current period (e.g., month);
- q represents the number of products considered;
- $j \in [1, q]$ represents the index for the current product;
- i_j represents the number of inventory count events for product j over periods $[1, T]$;
- $i \in [1, i_j]$ represents the index for the current count for product j ;
- $NetSales_j(i)$ represents the *recorded* cumulative sales quantities of product j , net of returns, from count $i - 1$ to count i ;
- $NetSales_j(i, t)$ represents the *book* cumulative sales quantities of product j , net of returns, between count $i - 1$ to count i that are allocated to period t ;
- $Discrepancies_j(i)$ represents the difference between the Book Inventory and the Physical Inventory observed at count i ;
- $Discrepancies_j(i, t)$ represents the difference between the Book Inventory and the Physical Inventory observed at count i allocated to period t .

Throughout the current section, the index j is dropped since only one product is being considered.

In this model, the following assumptions are made:

1. All demand transactions are unit-sized.
2. The probability of a demand transaction being recorded is $0 < p_t < 1$ and follows a Bernoulli distribution. It is constant within the period t and, therefore, is independent of demand.
3. No product returns can take place. This implies that *NetSales* is nonnegative.
4. All counts are perfectly accurate: they reflect the “true” physical inventory.

Assumption 1 allows us to equate the number of successes with the recorded sales quantities since the last count, and the number of failures with the discrepancy observed during the count.

The chosen prior for the Recording Accuracy p_t is a Beta distribution with the following hyper parameters:

$$\widehat{p}_t^0 \sim \text{Beta}(\alpha_t^0, \beta_t^0)$$

The Beta distribution is quite flexible and allows for a varying degree of confidence in the prior estimates. For instance, a weakly informative prior such as $\alpha_t^0 = 1$, $\beta_t^0 = 1$ reduces to the *Uniform(0,1)* distribution. At each count i , the estimate can easily be updated because the Beta distribution is a conjugate prior for this problem. In other words, the posterior (updated estimate) has a Beta distribution with the following hyper parameters:

$$\widehat{p}_t^{i+1} \sim \text{Beta}(\alpha_t^i + \text{NetSales}(i, t), \beta_t^i + \text{Discrepancies}(i, t))$$

where:

- $NetSales(i,t)$ represents the sum of recorded sales quantities, net of returns, since the last count that took place during period t . This quantity can be interpreted as the number of successes in a binomial experiment.
- $Discrepancies(i,t)$ represents the sum of unrecorded sales quantities, net of returns, since the last count that took place during period t . This quantity can be interpreted as the number of failures in a binomial experiment.

The equations below follow by induction:

$$\hat{p}_i \sim \text{Beta}(\alpha_i^0 + \sum_{l=1}^i NetSales(l,t), \beta_i^0 + \sum_{l=1}^i Discrepancies(l,t))$$

$$E[\hat{p}_i] = \frac{\alpha_i^0 + \sum_{l=1}^i NetSales(l,t)}{\alpha_i^0 + \sum_{l=1}^i NetSales(l,t) + \beta_i^0 + \sum_{l=1}^i Discrepancies(l,t)}$$

The confidence interval (or more accurately, the Bayesian credible interval) characterizing the uncertainty around the estimate is known. Note that the Maximum Likelihood Estimate of this Binomial proportion (i.e., simply taking the ratio of cumulative net sales to cumulative net sales and discrepancies) corresponds to the improper prior $\alpha_i^0 = 0, \beta_i^0 = 0$.

Allocation of count data to different periods

If the current and last count took place during period t , $Discrepancies(i,t)$ is readily available from the transactional data. However, if the time between two successive counts spans multiple periods, then only $NetSales(i,t)$ is known. $Discrepancies(i,t)$ are unknown since, by definition, there is no indication of when the inaccuracies were introduced. However, $Discrepancies(i)$, the difference between the book inventory record before the count and the true inventory after the physical count, is known.

In this case, $Discrepancies(i)$ are allocated to different periods according to a heuristic to produce estimates for the different periods. For instance, consider the case of a product for which estimates of monthly Recording Accuracy in June and July are sought. The physical inventory counts took place on the following dates (see Figure 5-5):

- $i = 1$: June 8th
- $i = 2$: July 14th
- $i = 3$: August 7th
- $i = 4$: September 9th

The proposed heuristic allocates discrepancies proportionally to the time spanned by the count in that period. For instance, 100% of the first count is allocated to the month of June because in this example no data are available prior to June 1st. The second count ($i = 2$) is allocated to June and July in the following proportions:

$$w_{t=June}^{i=2} = \frac{30 - 8}{14 + 30 - 8} = \frac{22}{36} = 61.1\%$$

$$w_{t=July}^{i=2} = \frac{14}{14 + 30 - 8} = \frac{14}{36} = 38.9\%$$

More generally, the following relationship defines the allocation heuristic:

- $Discrepancies(i,t) = Discrepancies(i) * w_t^i$, where w_t^i is equal to the percentage of the time since count $i - 1$ that took place during period t , such that the sum of the weights w_t^i over all periods is equal to one.

For computational convenience, instead of calculating $NetSales(i,t)$ from the transactional data, net sales are also allocated proportionally to the time spent in that period²¹:

- $NetSales(i,t) = NetSales(i) * w_t^i$, where $NetSales(i)$ is the sum of recorded sales quantities, net of returns, since the last count;

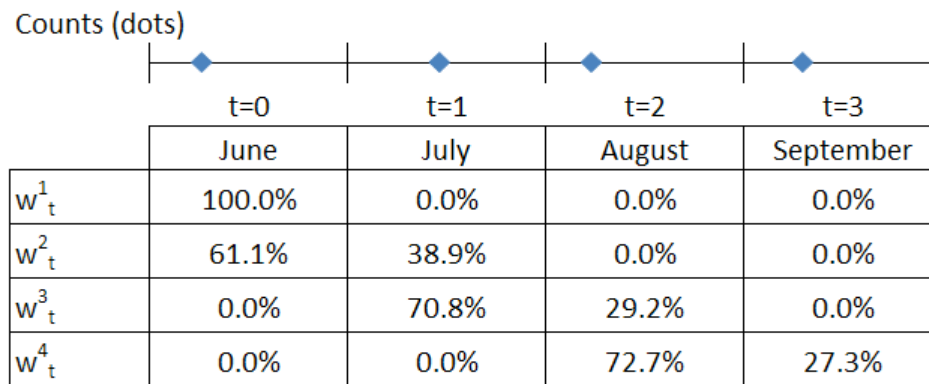


Figure 5-5 Count events (blue dots), periods and corresponding allocation weights

Note that although this approach will result in non-integer quantities, which may not make sense physically for some items, but which is not a problem with Bayesian estimation since the Beta distribution is defined over real positive parameters.

Bayesian credible interval

Given the calculated hyper parameters, the x% Bayesian credible interval is calculated by inverting the cumulative density function of the Beta distribution. As the parameters of the Beta distri-

²¹ Sensitivity analyses showed that the results did not differ materially.

bution increase, the width of such an interval decreases, reducing the uncertainty around the estimate of the Recording Accuracy.

5.4.3 Estimation for multiple products and with heterogeneity

In this subsection, all of the quantities described in 5.4.2 have an additional index j , which denotes a particular product and allows for heterogeneity in the Recording Accuracy that may occur for multiple reasons.

First, there could be unobserved differences between products that create variations in Recording Accuracy. For instance, products of different monetary value may exhibit radically different Recording Accuracy rates despite being stored in the same location. For instance, the staff may place more importance on properly recording transactions with such products. Anecdotally, it was noted that in a particular hospital studied, the total demand for batteries increased sharply before Christmas, because a large fraction went unrecorded, and this resulted in increased stock outs. Second, some measurement error in observations is likely, which may create overdispersion. For instance, a possible source of measurement error is incorrect counts. Lastly, by assuming no variation in Recording Accuracy, the previous method places proportionally heavier weights on fast-moving products at the expense of slow-moving products. For all of these reasons, it is important to allow for and to quantify heterogeneity in Recording Accuracy.

Skellam (1948) formally proposed the Beta-Binomial Model (BBM), which has been widely applied throughout the marketing literature and in ecology as a model for heterogeneity in a population. Lee and Sabavala (1987) and Ennis and Bi (1998) review the uses of the Beta-Binomial Model in different fields.

The Beta-Binomial Model (BBM) is obtained by assuming that a (fixed) Recording Accuracy $p_{\text{product}=j, \text{period}=t}$ for each item and period is sampled from a Beta distribution with parameters $\alpha_{\text{period}=t}$, $\beta_{\text{period}=t}$ (see Figure 5-6). Note that this differs from the case in the previous section: each

probability is considered fixed for a particular product and period, but the probabilities may differ among products. By estimating the parameters of the Beta distribution, the Recording Accuracy for the “average” product and the dispersion around this value among different products are characterized.

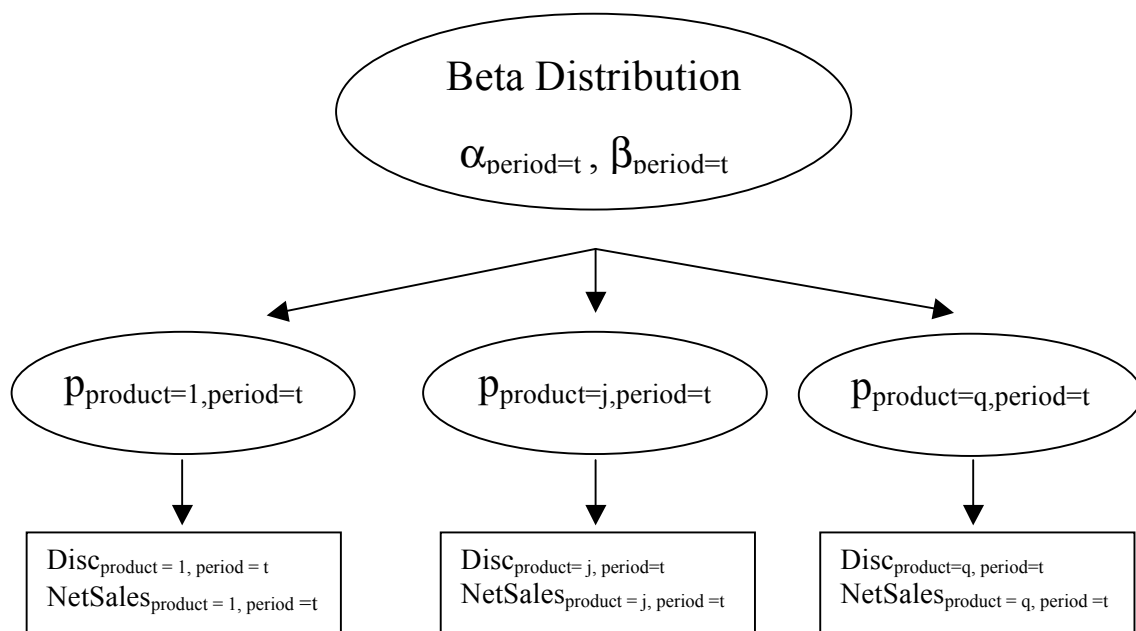


Figure 5-6 Hierarchical representation of the Beta-Binomial Model (BBM)

Difficulties in Bayesian Estimation of the Beta-Binomial Model

Bayesian estimation requires a prior assumption on the distributions of the hyper parameters α_t, β_t . Unfortunately and contrary to the Binomial case, no simple true conjugate priors exist for the Beta-Binomial distribution. Lee & Sabavala (1987) describe a prior with a conjugate-like property for the case when the number of trials is fixed and equal to two, which is too restrictive to be applied in the case of inventory inaccuracies. Everson & Bradlow (2002) derive closed-form expressions of the posterior distribution using polynomial expansions that are computation-

ally efficient, but their algorithm is quite complex, especially when the number of trials is not constant. Another alternative is to resort to numerical integration or Monte Carlo Markov Chain techniques. Because the benefit of Bayesian estimation of these quantities is limited in this case, and these approaches are computationally more expensive, Maximum Likelihood Estimation is used instead.

Maximum Likelihood Estimation of the Beta-Binomial Model

This paragraph outlines the procedure for Maximum Likelihood estimation. Writing the marginal likelihood function for an individual product observation j results in:

$$\begin{aligned}
 \mathcal{L}(\alpha_t, \beta_t \mid NS_j(t) = m, Disc_j(t) = n) &= P(NS_j(t) = m, Disc_j(t) = n \mid \alpha_t, \beta_t) \\
 &= \int P(NS_j(t) = m, Disc_j(t) = n \mid p_{j,t})P(p_{j,t} \mid \alpha_t, \beta_t)dp_{j,t} \\
 &= \int Binomial(m \mid y_j = m + n, p_{j,t})Beta(p_{j,t} \mid \alpha_t, \beta_t)dp_{j,t} \\
 &= \binom{m+n}{m} \frac{B(m + \alpha_t, n + \beta_t)}{B(\alpha_t, \beta_t)} \\
 &= \binom{NS_j(t) + Disc_j(t)}{NS_j(t)} \frac{B(NS_j(t) + \alpha_t, Disc_j(t) + \beta_t)}{B(\alpha_t, \beta_t)}
 \end{aligned}$$

where $Beta$ is the probability distribution function of the $Beta(\alpha_t, \beta_t)$ distribution and B is the Beta function.

Combining the observations for the different products j with $j \in [1, q]$:

$$\mathcal{L}(\alpha_t, \beta_t \mid Observations) = \prod_{j=1}^q \binom{NS_j(t) + Disc_j(t)}{NS_j(t)} \frac{B(NS_j(t) + \alpha_t, Disc_j(t) + \beta_t)}{B(\alpha_t, \beta_t)}$$

Taking the log and ignoring constants specific to the observations, the log-likelihood to maximize is:

$$L(\alpha_t, \beta_t | Observations) = -q \log(B(\alpha_t, \beta_t)) + \sum_{j=1}^q \log(B(NS_j(t) + \alpha_t, Disc_j(t) + \beta_t))$$

This expression can then be maximized using a numerical algorithm, as there is no known closed form solution. As discussed in Griffiths (1973), Williams (1975) and Prentice (1986), it is helpful to re-parameterize the Beta distribution to achieve faster convergence and numerical stability. The parameterization described in Prentice (1986) is applied:

$$p = \frac{\alpha}{\alpha + \beta}, \delta = \frac{1}{\alpha + \beta + 1}, \gamma = \frac{\delta}{1 - \delta}$$

The inverse formulas provide the original parameters of the Beta distribution:

$$\alpha = p\gamma^{-1}, \beta = (1 - p)\gamma^{-1}$$

Here p is the Recording Accuracy for the average product and δ is a positive coefficient that increases toward one with heterogeneity in the population.

5.5 Empirical Analysis of Recording Accuracy

In this section, the multi-item Beta Binomial Model is applied to transactional data from Lambda over a 15-month period.

5.5.1 Data pre-processing

Because of simplifying assumptions, real-world data contain some observations that are incompatible with the structure of the prior models. The treatment of such cases is discussed in this subsection. Because of the presence of product returns, the *NetSales* variable may take negative values. Given returns and other sources of inventory inaccuracies, instances where the *Discrepancies* variable takes negative values also occur (i.e., because discrepancies are defined as book

inventory minus physical inventory, this means that the physical inventory was understated). A frequency distribution of these instances for all count events is given in Table 5.2 below.

	Number of observations	% of observations
Counts	312,589	100.0%
Negative Discrepancies	14,339	4.6%
Negative Net Sales	3,912	1.3%
Negative Discrepancies AND Negative Net Sales	293	< 0.1%
Negative Discrepancies OR Negative Net Sales	17,958	5.7%

Table 5.2 Distribution of negative values among count events

Overall, negative observations account for 5.7% of observations. Such negative observations, in which Net Sales and/or Discrepancies are negative, are added to the Net Sales and Discrepancies values of the next count event in time, until non-negative values are obtained. This approach avoids bias by capturing the information, and retains the fact that a count took place for the purpose of calculating the time since the last count.

Because of the allocation heuristic described in 5.4.2, Net Sales and Discrepancies allocated to a given time period may take non-integer values. Both the Bayesian updating framework and Maximum Likelihood Methods assume discrete observations, and therefore this could potentially be problematic. In the case of the Bayesian updating model, the Beta distribution is defined for real positive parameters, and because of the simplicity of the formulas described in 5.4.2, it is clear that they extend smoothly to the case of non-integer observations. On the other hand, Maximum Likelihood Estimation requires discrete observations. To carry out the estimation of the Beta-Binomial Model, $NS_j(t)$ and $Disc_j(t)$ are rounded to the nearest integer.

5.5.2 Implementation

Transactional data from Lambda were exported into a MySQL database after allowing for coding to reduce the size of the database and for optimized processing. From the transactional data, the

quantities described in the outlined data are calculated using code programmed in the R statistical programming language and SQL queries to the database. The Maximum Likelihood Estimation is performed through an iterative numerical optimization routine provided in the R VGAM package (Yee, 2008).

5.5.3 Product groups

To improve the confidence intervals on the estimates of the Recording Accuracy, increasing the number of count event observations is desirable. One approach is to consider more products and to quantify the heterogeneity in the Recording Accuracy using the Beta-Binomial Model (BBM). However, the Beta distribution is inappropriate if the distribution of the Recording Accuracy is not unimodal. Therefore, products likely to have radically different recording accuracy should be isolated rather than pooled together. Furthermore, from a managerial perspective, identifying factors that account for different recording accuracy is important, as they may suggest possible improvements.

Since the Beta-Binomial Model outlined in 5.4.3 allows for and quantifies the heterogeneity in Recording Accuracy, it is actually not critical to immediately obtain the correct product segmentation. An ineffective segmentation will be reflected in the parameter estimates of the Beta distribution, which will reflect a high level of heterogeneity (i.e., values for α_i , β_i below or around one). A new segmentation can then be devised and the BBM model applied again until the degree of heterogeneity within each group of products is reduced to an acceptable level.

As a first step, groups (each having their distribution hyperparameters α_i , β_i) are defined according to stations. A possible refinement to this approach is to distinguish between products available behind closed cabinet doors and products available on open shelves (because accessibility could influence Recording Accuracy).

5.5.4 Results

This section applies the Beta-Binomial Model described in section 5.4.3 to analyze all products between January 1, 2007 and April 1, 2008. The Recording Accuracy is computed at monthly intervals ($T = 15$ months) and plotted for a subset of stations: 6C, 6E, 6SE, 5W, Cath Lab A, Cath Lab B and Cath Lab, and the median Recording Accuracy is estimated across all stations.

One can note in Figure 5-7 that the Cath Lab stations (CL A, B and C) have relatively low levels of Recording Accuracy (75–90%), although they have improved a bit over time. This was explained by the presence of a full-time technician in charge of performing frequent counts and re-ordering high-value supplies such as stents and catheters. A possible hypothesis is that the presence of this full-time employee created a “safety net” and therefore made it less critical for staff to actually record their inventory usage. The sharp drop in the median Recording Accuracy in March 2007 was caused by the installation of new machines with very few SKUs and exceptionally low recording accuracy during the training phase.

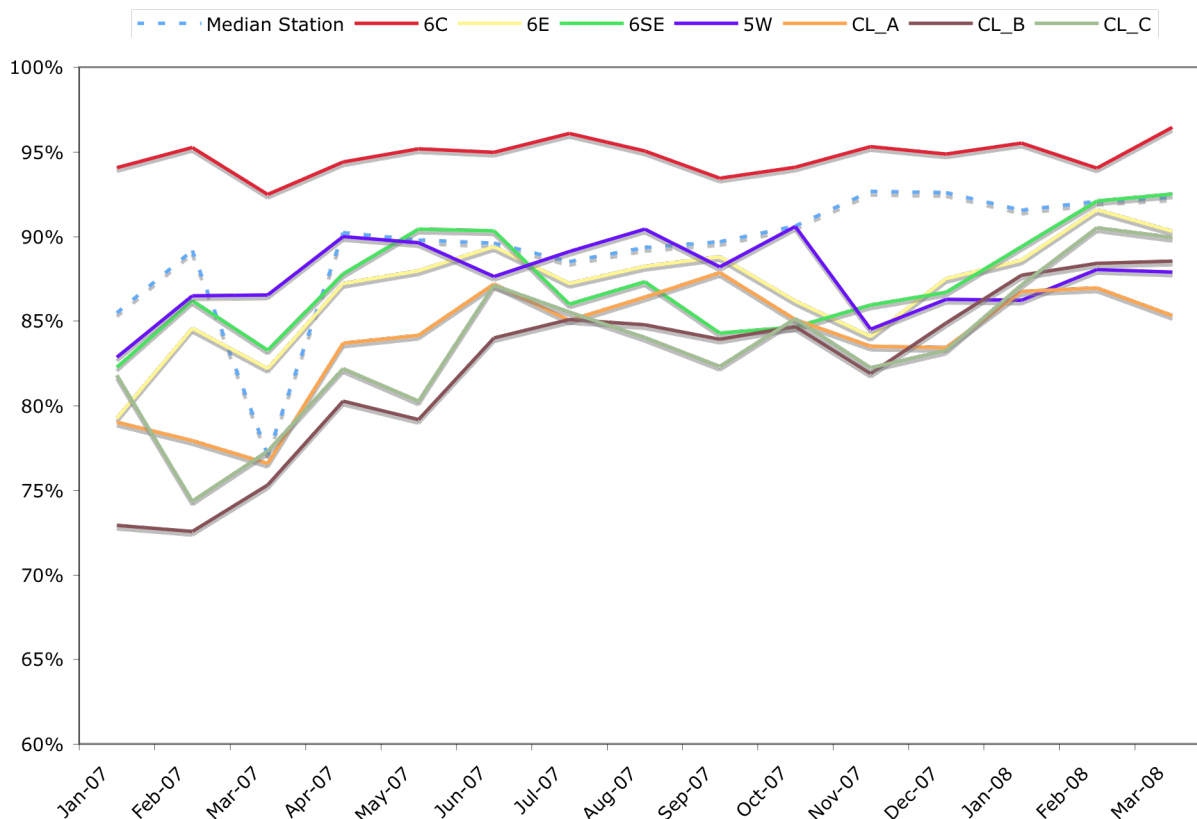


Figure 5-7 Beta-Binomial Model: Monthly estimates of the Recording Accuracy for seven stations

Lambda has a separate facility located approximately 50 miles North of the main campus of the hospital. ADMs were deployed at Lambda North and had higher Recording Accuracy than those in wards that performed the same services and operating rooms available at Lambda’s main facility, as shown in Figure 5-8. This suggests that high levels of Recording Accuracy are achievable. However, it is difficult to pinpoint the factor that led to this higher performance. Several hypotheses were advanced by hospital managers: senior nursing staff being more committed to the success of the Automated Dispensing Machines and the lack of a night shift at Lambda North would limit the number of less-well trained people accessing the machines. I observed a staff meeting and it was apparent that the satisfaction with the system by nursing and materials man-

agement was higher at Lambda North, but did not identify a causal effect explaining this higher Recording Accuracy.

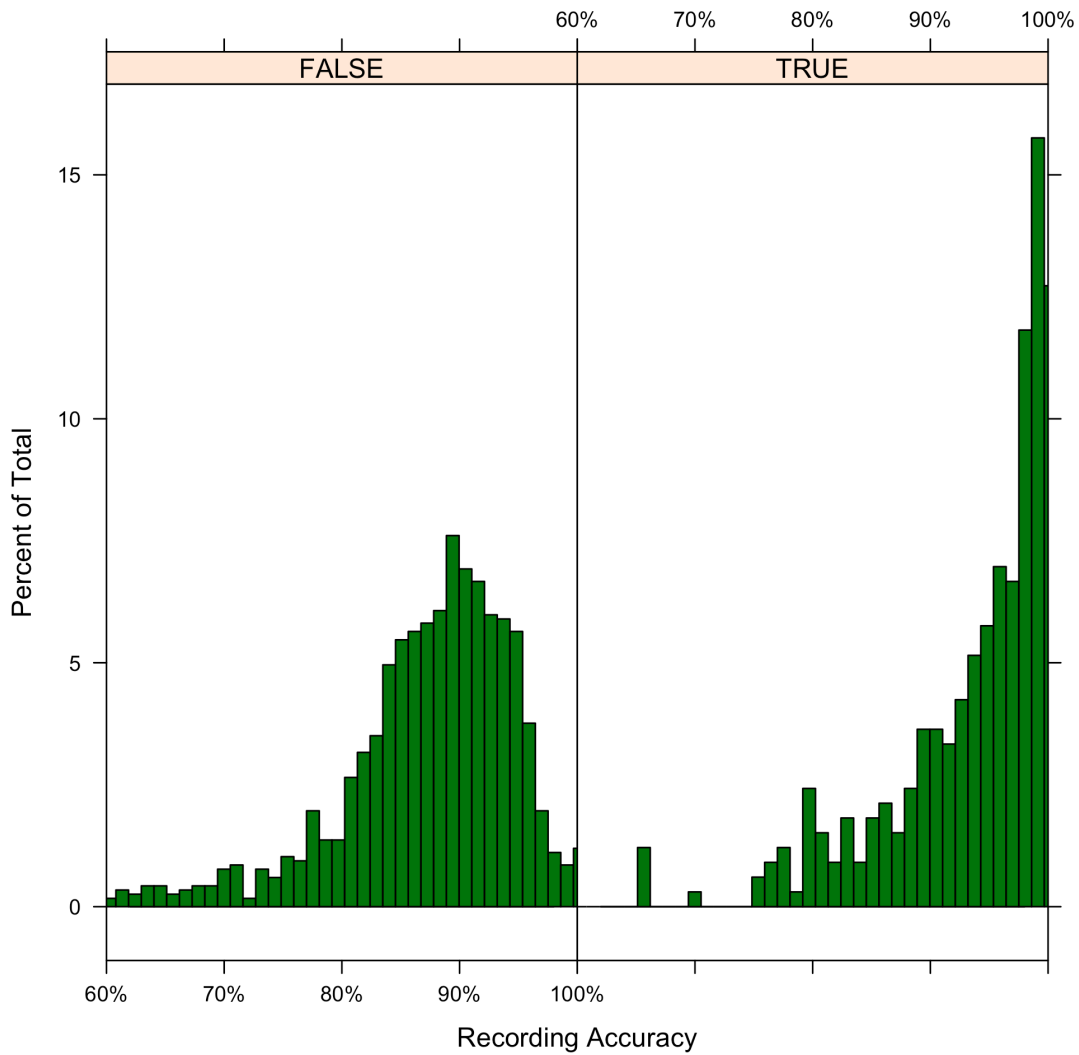


Figure 5-8: Histogram of Monthly Recording Accuracy, separating Lambda's Main facility (left pane) from Lambda North (right pane)

Figure 5-9 shows that 17 stations out of a total of 100²² had a median Recording Accuracy above 95%, 34 stations had a median Recording Accuracy between 90% and 95% and 49 stations had a Recording Accuracy below 90%. These figures indicate that Imperfect Demand Recording is large and an essential source of inventory inaccuracy at Lambda.

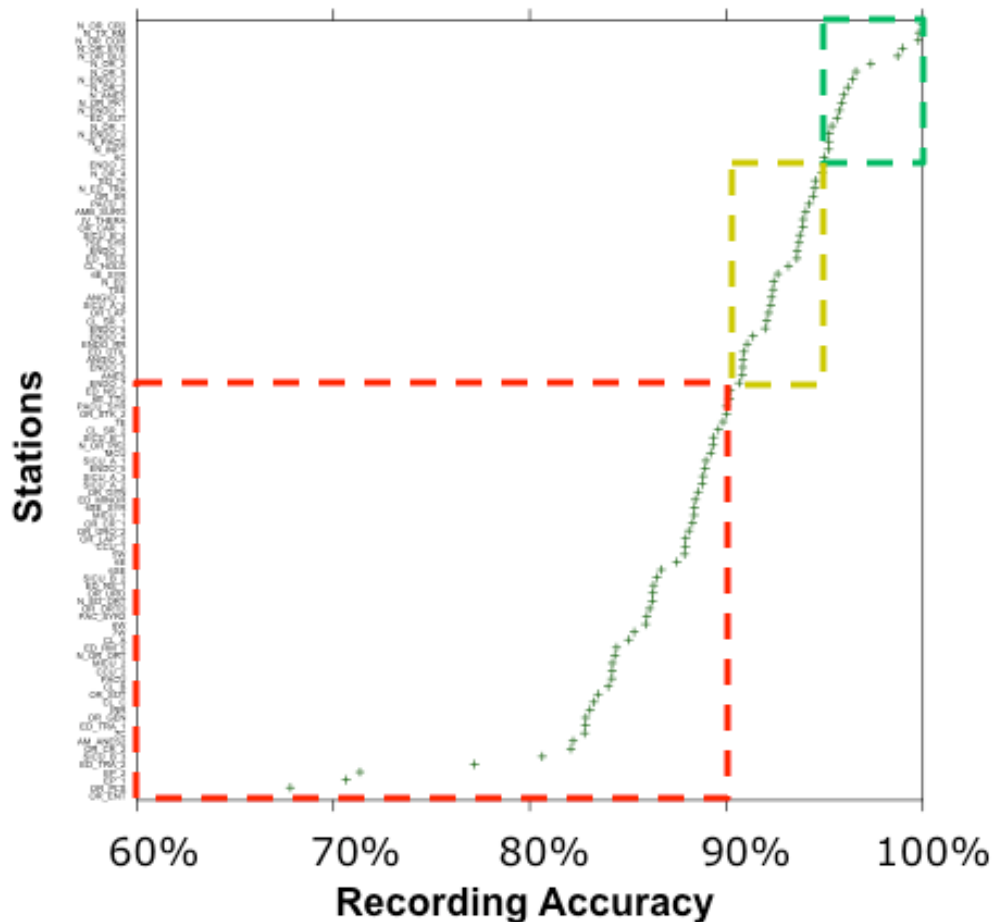


Figure 5-9: Stations ranked by Median Recording Accuracy across 15 months

²² Eight stations did not have sufficient counts to estimate Recording Accuracy during all months.

[Page intentionally left blank]

Chapter 6

Impact of Inventory Inaccuracies on System Performance

The goal of this chapter is to understand the effect of inventory inaccuracies on the performance of the hospital supply chain through modeling rather than data analysis. Because the analysis conducted in Chapter 5 concluded that Imperfect Demand Recording is the main source of inventory inaccuracies in the hospital inventory inaccuracies arising from this source are considered in more detail. Section 6.1 describes the research approach and a discrete-event simulation model of the phenomenon. Section 6.2 analyzes product availability under different scenarios, and section 6.3 studies the bias in performance metrics. The next chapter explores in more detail the effect of different counting policies.

6.1 Discrete-Event Simulation Model

Discrete-Event Simulation (DES) was chosen to reflect the stochastic nature of the inputs, and to provide flexibility in exploring the effect of different ordering and counting policies that are difficult to analyze in closed form (Riddalls, Bennett, & Tipi, 2000).

6.1.1 Key assumptions

This subsection summarizes the key assumptions of the model.

- Single-item model: only a single Station-SKU is considered. Interaction effects among products or locations, such as substitution effects during stock-outs, are very difficult to characterize and quantify from Automated Dispensing Machine data because of inventory inaccuracies and small sample sizes. Using a single-item model ignores these effects because the focus is on understanding the effects of inventory inaccuracies.
- Independent and identically distributed (i.i.d.) demand: while this is a common assumption in inventory models, checking the validity of this assumption in a hospital context is important. The presence of autocorrelation was tested by conducting a Ljung-Box test with lags up to 14 days²³, which failed to reject the null hypothesis of independence of recorded demand over time.
- Negative binomial hourly demand: among distributions appropriate for modeling count data, the negative binomial (which is a Gamma Poisson mixture) offers flexibility in modeling situations with a high number of zeros and/or variance greater than the mean. Agrawal & Smith (1996) have shown that retail demand is better approximated using a negative binomial distribution than a normal distribution, and can accurately describe both low-demand and high-demand items.
- Bernoulli recording: recording of demand is modeled by sampling from a Bernoulli distribution with a fixed parameter p_t , and therefore the recorded demand²⁴ is a Negative Binomial-Bernoulli mixture. In Appendix C, this mixture is shown to be equivalent to a negative binomial distribution with modified parameters.

²³ Our autocorrelation testing procedure is described in Appendix B.

²⁴ Assuming the product is in stock.

- (R, s, S) order review policy chosen to reflect the policy implemented by the ADMs: daily periodic review, reorder-point, and order-up-to policy. The details of the policy are explained below.
- All counts are perfect and reflect the true physical inventory²⁵.
- Time between counts is modeled either as a fixed parameter or sampled from a random variable, as follows:
 - o Fixed: every 7, 14, or 21 days
 - o Random Variable: sampled from a Gamma distribution (of which the exponential is a special case) or a Truncated Normal Distribution
- Counting on replenishment: this is implemented as a sampling from a Bernoulli random variable. The model is exercised with a probability of counting on replenishment equal to zero or one.

6.1.2 Model structure

The model maintains two inventory views in parallel: the “physical view” and the “book view.” The physical view indicates the inventory quantity physically available, and the book view indicates the inventory quantity recorded in the system, thus reflecting only recorded transactions.

The model is based on four key processes that represent logistical activities (see Figure 6-1). Every two hours, the Demand process is called and generates the number of units of demand by sampling from a negative binomial distribution. If this quantity is greater than zero, a demand transaction is defined and the physical view inventory level is decremented by the quantity with-

²⁵ This implies that imperfect counts could lead to lower service levels than those reported in this model.

drawn. Sampling from a Bernoulli distribution determines whether the demand transaction is recorded (i.e., typing errors in demand quantity are not considered). According to hospital data records, at least 70% of demand transactions were unit-sized, and this pattern was consistent across over 70% of SKUs, which suggests that this is a reasonable approximation for most products²⁶. A Bernoulli trial determines this binary outcome with probability equal to an input parameter representing recording accuracy. If the outcome is positive, the book view inventory level is decremented by the quantity withdrawn (for instance, five units if demand was five units and the five units were physically available). If the quantity demanded is greater than the physical inventory level, the demand will be only partially satisfied by whatever quantity is in the physical inventory level. That quantity, called “sales” or “usage,” is subtracted from the physical inventory level and may be subtracted from the book view inventory level if the demand transaction is recorded. Any demand that is not satisfied by the physical inventory available is recorded as “lost sales.”

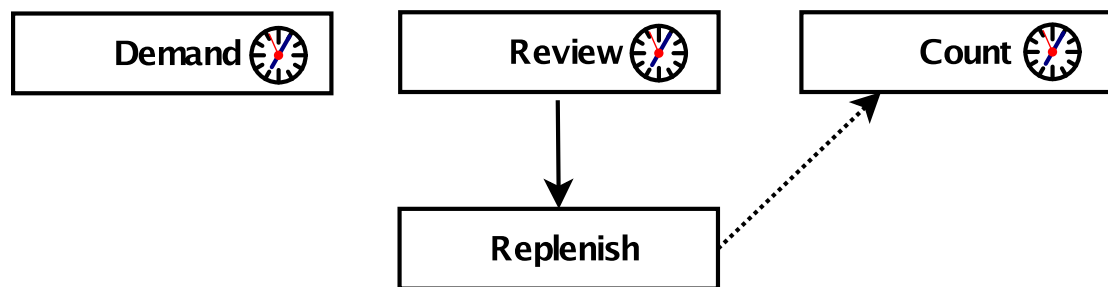


Figure 6-1: Key processes used in the simulation. The clock signifies periodicity.

²⁶ Again, the usage data from the hospital are biased, so it is possible that the number of unit-sized transactions is overestimated if end-users only press the button once when they take multiple products at a time. However, most products, by the nature of their function, are only used in unit-sized quantities on a given patient.

The Review process is called at specific times throughout the week, which in the base case corresponds to one review per day. Orders can only be placed during a Review. Consistent with the policy implemented in the hospital by the ADM software, orders are placed according to an (s,S) policy (i.e., an order is placed if the current book inventory level plus the sum of the order quantities of any unfilled orders is less than or equal to the reorder point, and the order size equals the order-up-to point minus the current book inventory level minus the sum of the order quantities of any unfilled orders). In this simulation, a fixed replenishment lead time is considered. Although the model is structured to allow for daily variation in time between reviews and replenishment lead times (for instance on weekends), it is exercised with constant time between reviews and replenishment lead times. This choice is motivated by a desire to keep the model relatively simple and conservative (i.e., to avoid compounding the effects of inventory inaccuracy with variance in lead-time²⁷).

After the lead-time has passed, the Replenishment process increments the physical and book view inventory levels by the order quantity. Furthermore, eighteen months after introducing ADMs to manage its medical and surgical process, the hospital started training replenishment technicians to perform counts prior to replenishing ADMs. This option of activating a Count process during calls to the Replenishment process is signified by the dashed line in Figure 6-1. The implications of this policy, which is called “Count on Replenishment,” are discussed in section 6.3.2 below. Furthermore, the model takes as input the probability of performing a count on a replenishment event, allowing the analyst to explore the effect of partial compliance with this policy by replenishment technicians.

²⁷ When we run the model based on a real hospital replenishment schedule, implying extended replenishment lead times on weekends, we find increased stock-outs frequencies and lower fill rates relative to this base case. This is not surprising as both the expected lead time and the variance of the lead time increase.

The Count process corresponds to the audit process and adjusts the book view inventory level to the physical view inventory level at the conclusion of each Count transaction. Perfect audits are assumed.

6.1.3 Implementation

In the model, events may occur at any time during the day, and the simulation uses a next-event time advance approach (Law, 2007). Although demand occurs at periodic intervals, it is a discrete-event model as opposed to a discrete-time model. There are two reasons for this choice: first, some stochastic variables such as the time between successive counts are better modeled as continuous rather than discrete variables. This choice does not lead to any loss of functionality, as discrete-time models can be viewed as a particular case of discrete-event models where events occur with fixed periodicity. Second, a discrete-event framework facilitates modeling at a higher level of abstraction: instead of thinking at a time-step level, the researcher is enabled to code only what happens at each process step, which is more intuitive and aligned with supply chain thinking. Implementing this paradigm in a high-level language leads to source code that is more compact and, therefore, easier to understand, debug and extend. The simulation is coded in Python using the SimPy open-source discrete-event simulation library (Muller, 2004), and the source code is presented in Appendix D.

Simulation outputs are written to a MySQL database and analyzed with the R statistical environment, as explained in (Team, 2004). For each alternative configuration, the model is exercised for 200 runs, each with a run length of 364 days (52 weeks). The run length was experimented with and increased until the means of different performance metrics approximately reached a steady state (Law, 2007). The number of runs was gradually increased to achieve sufficiently narrow confidence intervals on the difference between alternative configurations and the perfect recording case.

Consistent with the Common Random Numbers variance reduction technique, each random variable in the simulation has its own random stream, and a unique seed that depends only on the trial number. This synchronization ensures that a specific random number used for a specific purpose in one configuration is used for exactly the same purpose in all other configurations. Different performance metrics are computed over the entire length of the run (364 days) and averaged across trials, which provides estimates of the steady-state behavior of those metrics. Alternate configurations are compared using the nonparametric Wilcoxon signed-rank test. A nonparametric test is essential because many performance metrics (such as the fill rate or in-stock probability) are unlikely to follow a normal distribution²⁸. The Wilcoxon test applies to paired samples: each trial constitutes a pair of observations for the two configurations being compared. The test can be used to produce a confidence interval on the difference in the performance metric.

6.1.4 Base case

The base case represents a medium-to-fast moving medical/surgical item, such as a catheter for use in an operating room. The base case parameters are set based on Lambda's ADM data. The mean daily demand follows a negative binomial distribution, with a mean of one item per day and a variance-to-mean ratio of 1.667. There are no returns. Weekends are treated the same as weekdays. A Review is scheduled every day at 3 pm²⁹. When an order is placed during a Review, the shipment will arrive at the station 19 hours later, at 10 am the next day. This lead time reflects the proximity of the distributor's warehouse at a distance of less than 50 miles away

²⁸ Or any specific distributional assumption.

²⁹ See previous footnote.

from the hospital and the distributor's daily delivery policy. In the base case, all Replenishment transactions are recorded, and no counts take place upon replenishment. The reorder point is set to four units, the order-up-to level is set to ten units, and the initial inventory is set to seven units, consistent with the inventory policy parameters in place at the hospital for similar items. The effect of count periodicity and imperfect recording are explored through a full factorial design. The probability of accurately recording each demand takes values between 0.65 and 1.00, in 0.05 increments. Count transactions are performed with a constant count periodicity of 7, 14, or 21 days. This results in 24 alternative configurations being simulated. The sensitivity of the model was tested by exercising it at constant count periodicities of 4, 10, 28, 35, 42, or 56 days, resulting in a total of 72 configurations. The results did not differ qualitatively from the effects presented. The different input parameters and their respective base-case values are summarized in Table 6.1.

Recording Accuracy	65% – 100%
Count periodicity	7, 14, or 21 days
Initial inventory level	7 units
Reorder point	4 units
Order-up-to level	10 units
Review event periodicity	Daily, 3 pm daily
Lead time	19 hours (delivery at 10:00 am next day)
Probability of performing a count on replenishment event	0 or 1
Mean demand (Negative Binomial distribution)	1 unit / day
Variance to mean ratio (Negative Binomial distribution)	1.667 units
Simulation Run Length	364 days
Trials	200

Table 6.1: Summary of input parameters and base case values.

6.2 Effect of Imperfect Demand Recording on Service Level Metrics

This section seeks to understand the effect of imperfect recording on physical service level metrics. The next section investigates whether imperfect recording introduces bias in service level metrics calculated from “book” (and hence potentially inaccurate) records.

Multiple potential service level metrics are used to explore these hypotheses. To choose service level metrics for the point-of-use availability of supplies, it is important to keep in mind the context and specificities of the hospital environment. Because unfilled demand is typically not backlogged³⁰, the fill rate metric is particularly aligned with the experience of clinician end-users, who were considered “customers” of this particular supply chain. However, ADMs only record product withdrawals from the station and not actual demand, making the calculation of the fill rate from ADM data challenging. Agrawal & Smith (1996) propose a methodology for estimating the negative binomial demand in the presence of unobservable lost sales, which could be used to derive estimates of the fill rate. However, imperfect demand recording adds another layer of uncertainty and limits the applicability of this methodology. The average in-stock probability over the course of a simulation run is used as a substitute and corresponds to the probability that the product is in stock at a random point in time. This metric reflects the probability that a nurse will have to perform additional tasks to obtain the product. Because of this and the fact that, in the absence of inventory inaccuracies, in-stock probability can be computed from ADM data, it is a potential candidate for a level-of-service (LOS) metric. Another potential candidate is stock-out frequency, which Lambda is currently using; it is reported by the Automated Dispensing Machine information system. This LOS metric is defined as the number of stock-out events (defined as instances when the book inventory level reaches zero) per period. A period is usually a week or a month. While stock-out frequency is imperfect in that it does not account for the duration of stock-out events, and therefore does not reflect the end-user experience, it is the only LOS metric available to hospital inventory managers, and therefore it is investigated in this study.

³⁰ The question of what happens during stock-outs is highly dependent on the medical context generating the product need, as well as product categories and local hospital policies. Depending on the situation, the product may be obtained from another hospital ward, from a nearby hospital, express-shipped from the vendor, a substitute may be used, or more rarely, a procedure may be rescheduled.

By definition, the Type I service level is equal to the probability of stocking out during a replenishment cycle. Therefore, it is equal to the average stock-out frequency divided by the average replenishment frequency, and can be estimated through simulation. Because its value depends on the replenishment frequency, it is not comparable across products with different replenishment frequencies. Moreover, for a single product, imperfect demand recording delays the triggering of replenishments and lowers the replenishment frequency, making the Type I service level metric less valid. It is considered here only for the purpose of comparing the results in this paper with that of existing literature.

Subsection 6.2.1 reports the effect of imperfect recording on physical in-stock probability; the effect on the physical stock-out frequency is reported in 6.2.2. The results of this analysis are compared with those of Morey (1985) in 6.2.3, and the effect of variance in count periodicity is discussed in 6.2.4.

6.2.1 Fill rate and physical in-stock probability

When imperfect recording is introduced, discrepancies arise between physical view and book view inventory levels. Book view inventory is always greater than or equal to physical view inventory because it is assumed that there are no returns, that replenishments are always recorded and that imperfect demand recording is the only source of inventory inaccuracies (other sources are described and evaluated in section 5.1).

As Recording Accuracy increases, in-stock probability increases, as shown in Figure 6-2, as well as the fill rate (Table 6.2). This effect is not only statistically significant, as shown in Table 6.2, it also has practical significance. For example, at a count periodicity of 14 days, increasing the recording accuracy from 80% to 100% increases the in-stock probability from 91% to 98.2%. The size of this effect is quantified by constructing confidence intervals on the difference in in-stock probability between the current configuration and the configuration corresponding to per-

fect recording accuracy using the nonparametric paired Wilcoxon signed-rank test (each test was conducted on 2 * 200 trial observations). These bounds are also reported in Table 6.2.

This analysis suggests that unless mitigated by frequent inventory counts, lack of recording accuracy has severe detrimental effects on service levels, as measured by in-stock probability.

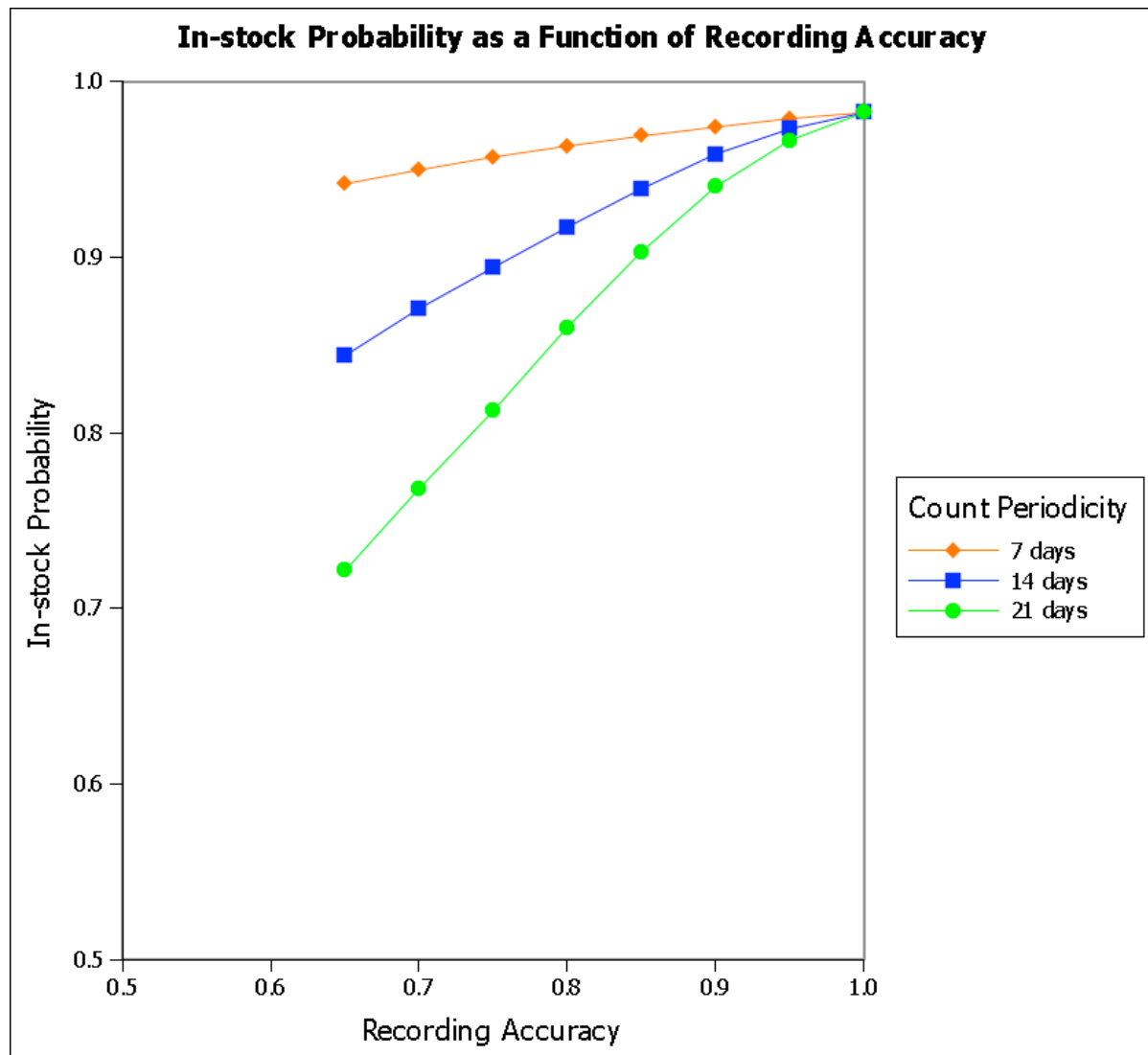


Figure 6-2: Effect of Count Periodicity on In-Stock Probability

Time Between Counts (days)	Recording Accuracy	Fill Rate	Physical In-stock Probability	95% Confidence Interval on the In-stock Probability Difference with the Perfect Recording Accuracy case		p-value
				Lower Bound	Upper Bound	
7	0.650	0.935 (0.024)	0.942 (0.017)	-0.042	-0.038	<0.0001
	0.700	0.944 (0.022)	0.950 (0.016)	-0.034	-0.030	<0.0001
	0.750	0.952 (0.019)	0.957 (0.013)	-0.027	-0.023	<0.0001
	0.800	0.958 (0.017)	0.963 (0.012)	-0.020	-0.017	<0.0001
	0.850	0.965 (0.015)	0.969 (0.010)	-0.014	-0.011	<0.0001
	0.900	0.971 (0.014)	0.974 (0.009)	-0.009	-0.007	<0.0001
	0.950	0.978 (0.012)	0.979 (0.007)	-0.004	-0.003	<0.0001
	1.000	0.982 (0.011)	0.982 (0.005)	NA	NA	NA
14	0.650	0.834 (0.042)	0.844 (0.039)	-0.142	-0.132	<0.0001
	0.700	0.862 (0.039)	0.871 (0.035)	-0.115	-0.105	<0.0001
	0.750	0.886 (0.037)	0.894 (0.030)	-0.092	-0.083	<0.0001
	0.800	0.908 (0.033)	0.917 (0.027)	-0.068	-0.060	<0.0001
	0.850	0.932 (0.028)	0.939 (0.023)	-0.045	-0.039	<0.0001
	0.900	0.954 (0.023)	0.959 (0.018)	-0.024	-0.020	<0.0001
	0.950	0.971 (0.016)	0.973 (0.011)	-0.009	-0.006	<0.0001
	1.000	0.982 (0.011)	0.982 (0.005)	NA	NA	NA
21	0.650	0.712 (0.054)	0.721 (0.056)	-0.268	-0.252	<0.0001
	0.700	0.759 (0.051)	0.768 (0.056)	-0.221	-0.205	<0.0001
	0.750	0.805 (0.050)	0.812 (0.050)	-0.176	-0.162	<0.0001
	0.800	0.850 (0.045)	0.860 (0.044)	-0.127	-0.114	<0.0001
	0.850	0.894 (0.041)	0.903 (0.037)	-0.082	-0.072	<0.0001
	0.900	0.934 (0.032)	0.940 (0.028)	-0.043	-0.036	<0.0001
	0.950	0.964 (0.020)	0.966 (0.016)	-0.016	-0.011	<0.0001
	1.000	0.982 (0.011)	0.982 (0.005)	NA	NA	NA

Table 6.2: Fill Rate and Physical In-stock Probability for different values of the Recording Accuracy and Count Periodicity³¹.

³¹ Standard errors in parentheses. 95% confidence intervals estimate the service level degradation due to Imperfect Demand Recording.

6.2.2 Physical stock-out frequency

A similar effect on the stock-out frequency is observed in Figure 6-4, which depicts the physical stock-out frequency as a function of recording accuracy and count periodicity: the physical stock-out frequency increases as the recording accuracy decreases.

Another observation can be made from the simulation results. The Physical Stock-out Frequency metric is problematic because it does not vary strictly monotonically with fill rate, in-stock probability or recording accuracy. For instance, for a count periodicity of 21 days, while recording accuracy decreases from 70% to 65%, physical stock-out frequency stays virtually unchanged at 0.403 stock-outs / week (Table 6.3). This is because additional stock-out events cannot take place once the physical inventory has already reached zero. Stock-out events are defined here as the time during which the on-hand inventory becomes equal to zero³². Meanwhile, the fill rate and in-stock probability are in fact deteriorating as a result of increasing stock-out duration. This increase in stock-out duration is because of the late detection of stock-outs, meaning that inventory managers should avoid relying on this metric when it is possible that stock-out durations are large, for instance for slow-moving products with low safety stocks.

³² However, from the perspective of an end-user, each time he/she goes to a station to get a unit of a particular SKU and this SKU is not available, this is perceived as a stock-out.

Time Between Counts (days)	Recording Accuracy	Physical View Stock-out Frequency (# stock-outs / week)	95% Confidence Interval on the Physical View Stock-out Frequency Difference with the Perfect Recording Accuracy case		p-value
			Lower Bound	Upper Bound	
7	0.650	0.281 (0.065)	0.135	0.154	<0.0001
	0.700	0.263 (0.069)	0.125	0.135	<0.0001
	0.750	0.243 (0.063)	0.096	0.115	<0.0001
	0.800	0.220 (0.060)	0.077	0.096	<0.0001
	0.850	0.197 (0.056)	0.058	0.077	<0.0001
	0.900	0.176 (0.051)	0.038	0.048	<0.0001
	0.950	0.155 (0.048)	0.019	0.029	<0.0001
	1.000	0.135 (0.046)	NA	NA	NA
14	0.650	0.391 (0.059)	0.250	0.269	<0.0001
	0.700	0.367 (0.067)	0.221	0.240	<0.0001
	0.750	0.343 (0.063)	0.202	0.212	<0.0001
	0.800	0.304 (0.061)	0.154	0.173	<0.0001
	0.850	0.267 (0.056)	0.125	0.135	<0.0001
	0.900	0.223 (0.061)	0.077	0.096	<0.0001
	0.950	0.177 (0.056)	0.038	0.048	<0.0001
	1.000	0.135 (0.046)	NA	NA	NA
21	0.650	0.404 (0.057)	0.260	0.279	<0.0001
	0.700	0.403 (0.060)	0.260	0.279	<0.0001
	0.750	0.384 (0.064)	0.240	0.260	<0.0001
	0.800	0.359 (0.067)	0.212	0.231	<0.0001
	0.850	0.321 (0.061)	0.173	0.192	<0.0001
	0.900	0.269 (0.063)	0.125	0.144	<0.0001
	0.950	0.200 (0.058)	0.067	0.077	<0.0001
	1.000	0.135 (0.046)	NA	NA	NA

Table 6.3: Physical View Stock-out Frequency (# stock-outs / week) for different values of the Recording Accuracy and Count Periodicity³³.

³³ Standard errors in parentheses. 95% confidence intervals estimate the service level degradation due to Imperfect Demand Recording.

6.2.3 Type I service level

Type I service level is obtained from the simulation by dividing the Physical Stock-out Frequency by the Replenishment Frequency for every simulation run, and then averaging this value across simulation runs, as displayed in Table 6.4.

This average value is compared with the value obtained for the corresponding set of parameters using the analytical formula in Morey (1985). The simulation model results in Type I service levels, which are 9%–22% lower than estimated using the Morey (1985) model, depending on the count periodicity and recording accuracy. Even with no inventory inaccuracies, the simulation model results in a Type I service level that is 12% lower than the Morey (1985) model, suggesting that this difference is at least partially explained by the choice of a negative binomial demand distribution. The Type I Service Level is the percentile of the lead time demand distribution corresponding to a value equal to the reorder point. For low demand items, a negative binomial distribution will have a lower corresponding percentile relative to a normal distribution of identical mean and variance. Another factor that may contribute to the gap in the presence of imperfect recording is that in the model physical inventory can never exceed book inventory (i.e., inventory errors can never improve the service level by adding excess inventory, whereas Morey's model assumes that the distribution of errors is symmetric).

Overall, this analysis shows that while Morey's model strives to be conservative through the use of a lower bound, in a hospital context the effects of imperfect recording may be more severe than previously suggested.

Count Periodicity	Recording Accuracy	Physical Stock-out Frequency (# stock-outs / week)	Replenishment Frequency (# repl. / week)	Simulated Type I Service Level	Analytical Type I Service Level	95% Confidence Interval on the difference between the simulated and analytical service level	
						Delta (min)	Delta (max)
7	0.650	0.281 (0.065)	0.913 (0.051)	0.692 (0.070)	0.799	-0.117	-0.097
	0.700	0.263 (0.069)	0.934 (0.050)	0.718 (0.072)	0.823	-0.116	-0.094
	0.750	0.243 (0.063)	0.952 (0.053)	0.745 (0.064)	0.849	-0.113	-0.094
	0.800	0.220 (0.060)	0.970 (0.052)	0.774 (0.059)	0.876	-0.111	-0.094
	0.850	0.197 (0.056)	0.983 (0.053)	0.800 (0.056)	0.905	-0.114	-0.097
	0.900	0.176 (0.051)	0.994 (0.054)	0.823 (0.050)	0.935	-0.120	-0.106
	0.950	0.155 (0.048)	1.003 (0.058)	0.846 (0.047)	0.965	-0.125	-0.112
	1.000	0.135 (0.046)	1.008 (0.059)	0.866 (0.044)	0.990	-0.131	-0.118
14	0.650	0.391 (0.059)	0.771 (0.054)	0.491 (0.083)	0.671	-0.191	-0.169
	0.700	0.367 (0.067)	0.819 (0.053)	0.551 (0.083)	0.702	-0.163	-0.139
	0.750	0.343 (0.063)	0.861 (0.056)	0.601 (0.073)	0.737	-0.146	-0.126
	0.800	0.304 (0.061)	0.901 (0.057)	0.663 (0.067)	0.777	-0.126	-0.107
	0.850	0.267 (0.056)	0.941 (0.057)	0.716 (0.059)	0.823	-0.116	-0.100
	0.900	0.223 (0.061)	0.973 (0.055)	0.771 (0.061)	0.876	-0.115	-0.097
	0.950	0.177 (0.056)	0.996 (0.058)	0.822 (0.055)	0.935	-0.121	-0.105
	1.000	0.135 (0.046)	1.008 (0.059)	0.866 (0.044)	0.990	-0.131	-0.118
21	0.650	0.404 (0.057)	0.639 (0.061)	0.365 (0.084)	0.589	-0.238	-0.213
	0.700	0.403 (0.060)	0.703 (0.062)	0.423 (0.085)	0.620	-0.210	-0.186
	0.750	0.384 (0.064)	0.769 (0.063)	0.498 (0.083)	0.657	-0.171	-0.148
	0.800	0.359 (0.067)	0.832 (0.062)	0.568 (0.083)	0.702	-0.144	-0.121
	0.850	0.321 (0.061)	0.895 (0.059)	0.641 (0.067)	0.756	-0.125	-0.105
	0.900	0.269 (0.063)	0.950 (0.059)	0.716 (0.066)	0.823	-0.117	-0.097
	0.950	0.200 (0.058)	0.988 (0.059)	0.797 (0.057)	0.905	-0.116	-0.100
	1.000	0.135 (0.046)	1.008 (0.059)	0.866 (0.044)	0.990	-0.131	-0.118

Table 6.4: Simulated Type I Service Level (Probability of stock-out in a replenishment cycle) for different values of the Recording Accuracy and Count Periodicity, as well as predicted Type I Service Levels according to Morey (1985) model.

6.2.4 Sensitivity to variability in Count Periodicity

The previous analyses reflect the fact that the service level decreases with increasing count periodicity, i.e., the average length of time between count events. When count periodicity is low, inventory discrepancies are promptly corrected and recording inaccuracies have a minor impact on service level (Figure 6-3). The base case assumes that counts are equally spaced in time (e.g., they happen *exactly* every 7 days). However, in practice, count events do not occur exactly at the scheduled time, and variations in the time between count events occur. Therefore, the entire distribution of the times between count events can affect the service level. The simulation models the variance in inter-count times using a Gamma distribution. The results show that as the variance of the inter-count times increases, the average service level decreases. For instance, if the inter-count times follow a Gamma distribution with a mean equal to 21 days and a coefficient of variation equal to 1.0 (i.e., an exponential distribution), with Recording Accuracy equal to 90%, the physical in-stock probability decreases from 94% to 84% solely as a result of the variability in inter-count times (Figure 6-3).

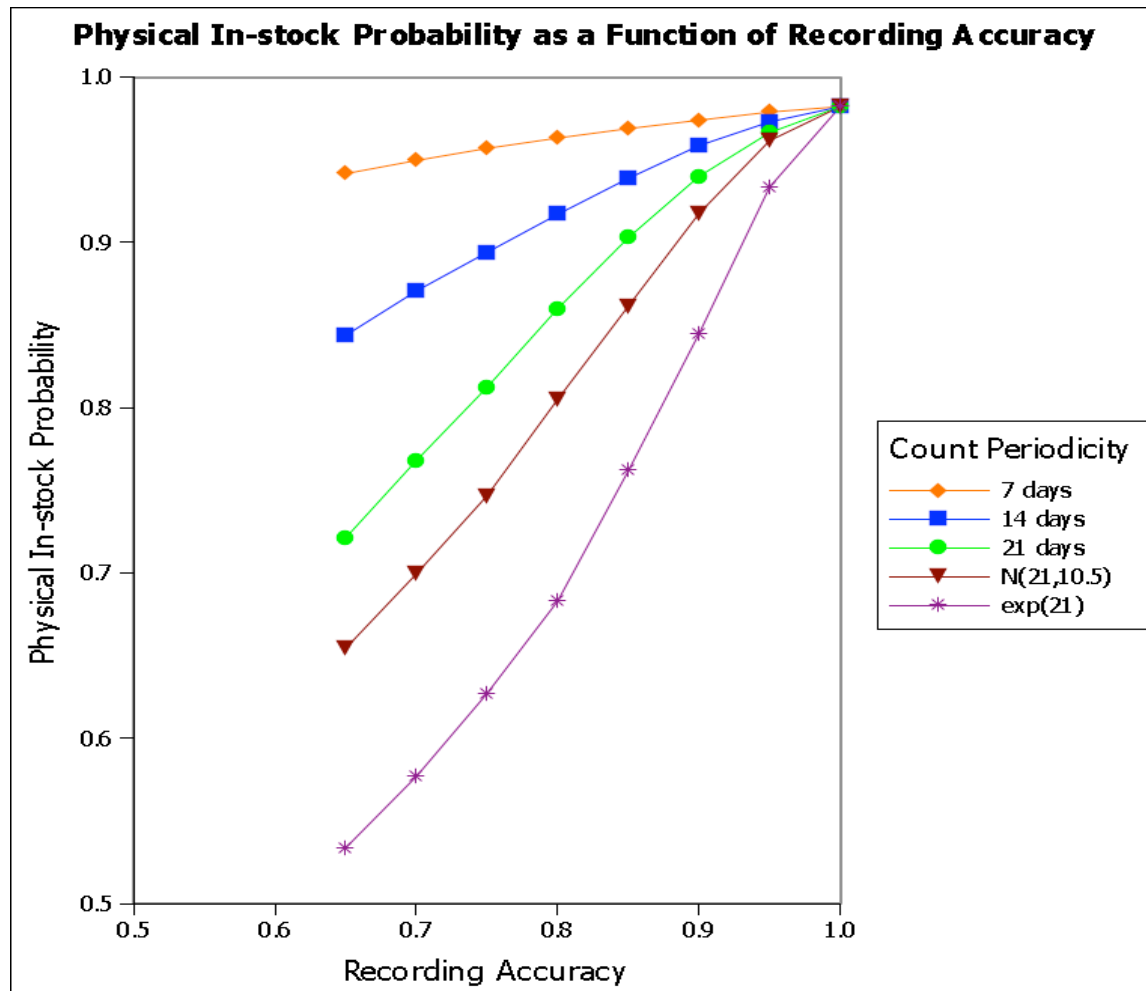


Figure 6-3: Effect of Variability of Inter-Count Times on In-Stock Probability, with a count periodicity of 21 days, a truncated normal distribution (C.V. = 0.5) and an exponential distribution (C.V. = 1)

Note: the parameters of the truncated normal were adjusted to ensure that the mean was equal to 21 days and the coefficient of variation was equal to 0.5.

6.3 Investigating the Bias in Book View Metrics

This section focuses on the effect of Recording Accuracy on the “book view” service level metrics, which are calculated based on data from the ADM inventory system. Because these data are inaccurate and overstate inventory levels under imperfect recording, it is likely that these metrics are biased. By maintaining parallel “physical” and “book” inventory records, the simulation framework allows us to test this hypothesis and quantify this bias.

6.3.1 Bias in the book view in-stock probability

Table 6.5 shows how the “book view” in-stock probability varies under different scenarios, and displays a confidence interval on the bias of this metric relative to the “physical” in-stock probability. The analysis shows that the physical in-stock probability can degrade materially while the book in-stock probability shows only a modest decline. For instance, at a count periodicity of 14 days and a recording accuracy of 65%, the physical in-stock probability is 84.4% and the book in-stock probability is 96.9%, implying a 12.5% bias. As shown by the confidence intervals of Table 6.5, this bias effect is statistically and materially significant, and increases as the physical in-stock probability decreases. Therefore, the book in-stock probability is severely misleading and will cause inventory managers relying on it to overestimate the service levels provided to end-users when recording accuracy is imperfect.

Count Periodicity	Recording Accuracy	Physical In-stock Probability	Book In-stock Probability	95% Confidence Interval on the metric bias	
				Delta (min)	Delta (max)
7	0.650	0.942 (0.017)	0.978 (0.006)	0.034	0.038
	0.700	0.950 (0.016)	0.980 (0.006)	0.029	0.032
	0.750	0.957 (0.013)	0.982 (0.005)	0.023	0.026
	0.800	0.963 (0.012)	0.983 (0.005)	0.019	0.021
	0.850	0.969 (0.010)	0.984 (0.005)	0.014	0.016
	0.900	0.974 (0.009)	0.984 (0.005)	0.009	0.010
	0.950	0.979 (0.007)	0.983 (0.005)	0.004	0.005
	1.000	0.982 (0.005)	0.982 (0.005)	0.000	0.000
14	0.650	0.844 (0.039)	0.969 (0.007)	0.119	0.129
	0.700	0.871 (0.035)	0.973 (0.006)	0.097	0.106
	0.750	0.894 (0.030)	0.977 (0.006)	0.078	0.086
	0.800	0.917 (0.027)	0.980 (0.005)	0.059	0.066
	0.850	0.939 (0.023)	0.983 (0.005)	0.041	0.047
	0.900	0.959 (0.018)	0.985 (0.005)	0.022	0.027
	0.950	0.973 (0.011)	0.984 (0.005)	0.009	0.011
	1.000	0.982 (0.005)	0.982 (0.005)	0.000	0.000
21	0.650	0.721 (0.056)	0.965 (0.005)	0.235	0.250
	0.700	0.768 (0.056)	0.969 (0.005)	0.192	0.207
	0.750	0.812 (0.050)	0.973 (0.005)	0.153	0.167
	0.800	0.860 (0.044)	0.978 (0.005)	0.111	0.122
	0.850	0.903 (0.037)	0.982 (0.005)	0.072	0.081
	0.900	0.940 (0.028)	0.984 (0.005)	0.038	0.046
	0.950	0.966 (0.016)	0.985 (0.005)	0.015	0.018
	1.000	0.982 (0.005)	0.982 (0.005)	0.000	0.000

Table 6.5: Physical and Book View In-stock Probability for different values of the Recording Accuracy and Count Periodicity

6.3.2 Bias in the book view stock-out frequency

The analysis shows that under periodic counts, the book view underestimates stock-out events and, therefore, also introduces significant bias in the stock-out frequency metric.

Table 6.6 shows an example of this bias at a count periodicity of 14 days. If replenishment occurs after a physical stock-out event but before the book view inventory level reaches zero (either through recorded demand or a count), the book view inventory record is increased, and a stock-out event has been missed.

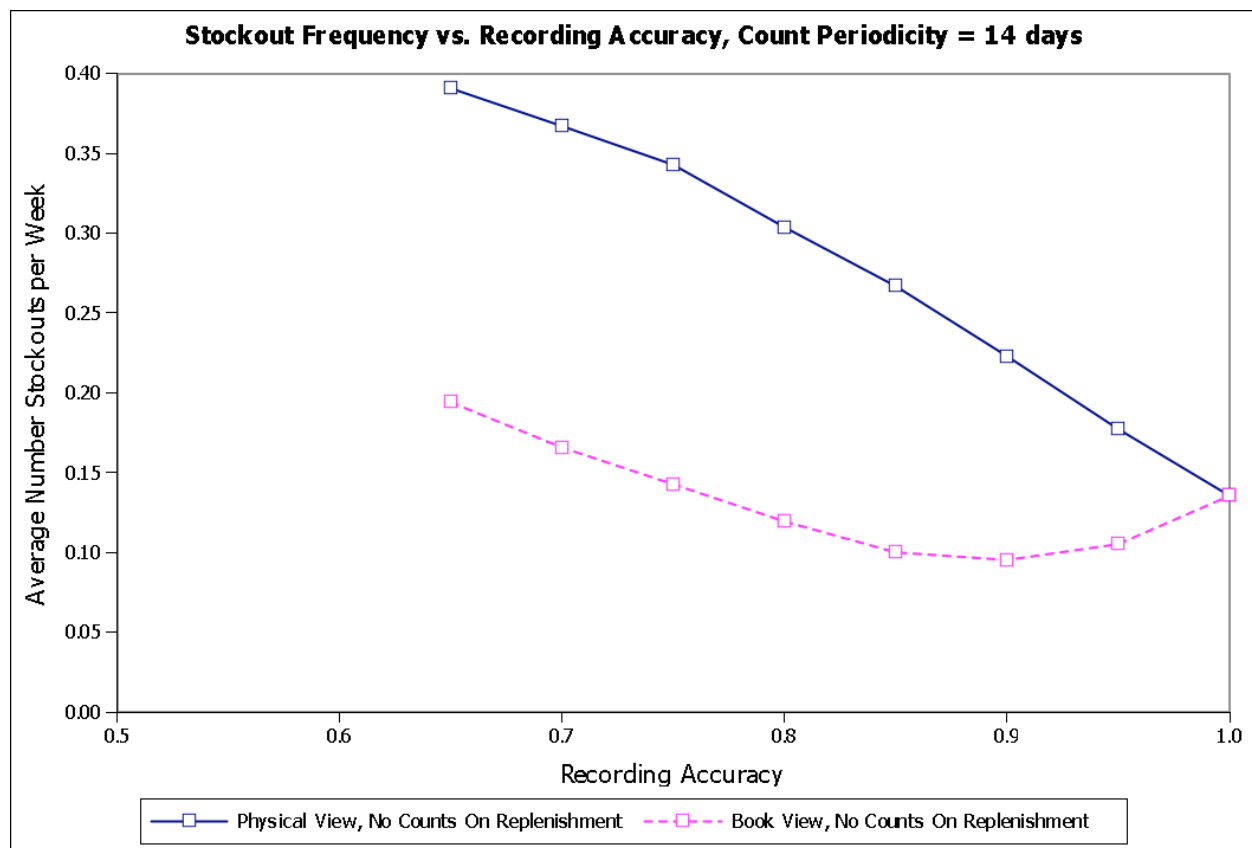


Figure 6-4: Physical and Book Stock-out Frequencies as a function of Recording Accuracy, at a Count Periodicity of 14 days.

				95% Confidence Interval on the metric bias	
Count Periodicity	Recording Accuracy	Physical View Stock-out Frequency (# stock-outs / week)	Book View Stock-out Frequency (# stock-outs / week)	Delta (min)	Delta (max)
7	0.650	0.281 (0.065)	0.136 (0.047)	-0.154	-0.135
	0.700	0.263 (0.069)	0.125 (0.047)	-0.144	-0.135
	0.750	0.243 (0.063)	0.116 (0.046)	-0.135	-0.115
	0.800	0.220 (0.060)	0.106 (0.043)	-0.115	-0.106
	0.850	0.197 (0.056)	0.103 (0.043)	-0.096	-0.087
	0.900	0.176 (0.051)	0.108 (0.042)	-0.067	-0.067
	0.950	0.155 (0.048)	0.120 (0.043)	-0.038	-0.038
	1.000	0.135 (0.046)	0.135 (0.046)	0.000	0.000
14	0.650	0.391 (0.059)	0.194 (0.048)	-0.202	-0.192
	0.700	0.367 (0.067)	0.166 (0.045)	-0.202	-0.192
	0.750	0.343 (0.063)	0.143 (0.042)	-0.202	-0.192
	0.800	0.304 (0.061)	0.119 (0.042)	-0.192	-0.173
	0.850	0.267 (0.056)	0.100 (0.039)	-0.173	-0.154
	0.900	0.223 (0.061)	0.095 (0.041)	-0.135	-0.125
	0.950	0.177 (0.056)	0.105 (0.044)	-0.077	-0.067
	1.000	0.135 (0.046)	0.135 (0.046)	0.000	0.000
21	0.650	0.404 (0.057)	0.223 (0.035)	-0.192	-0.173
	0.700	0.403 (0.060)	0.196 (0.040)	-0.212	-0.192
	0.750	0.384 (0.064)	0.165 (0.036)	-0.231	-0.212
	0.800	0.359 (0.067)	0.136 (0.039)	-0.231	-0.212
	0.850	0.321 (0.061)	0.110 (0.038)	-0.221	-0.202
	0.900	0.269 (0.063)	0.095 (0.039)	-0.183	-0.163
	0.950	0.200 (0.058)	0.098 (0.038)	-0.106	-0.096
	1.000	0.135 (0.046)	0.135 (0.046)	0.000	0.000

Table 6.6: Physical and Book View Stock-out Frequency (# stock-outs / week) for different values of the Recording Accuracy and Count Periodicity

Furthermore, as shown in Figure 6-4, the book stock-out frequency is not a monotonic function of Recording Accuracy. For instance, at a count periodicity of 14 days, for high Recording Accu-

racy (e.g., 95%) the book stock-out frequency is low and close to the physical stock-out frequency. As the Recording Accuracy level decreases toward 90%, the book stock-out frequency decreases toward its minimum. As the Recording Accuracy decreases further below 90%, the book stock-out frequency increases again. This result is robust across different count periodicities, as shown in Figure 6-5 and Table 6.6, and explained further below.

Stock-Out Detection Rate vs. Recording Accuracy (No Counts on Replenishment)

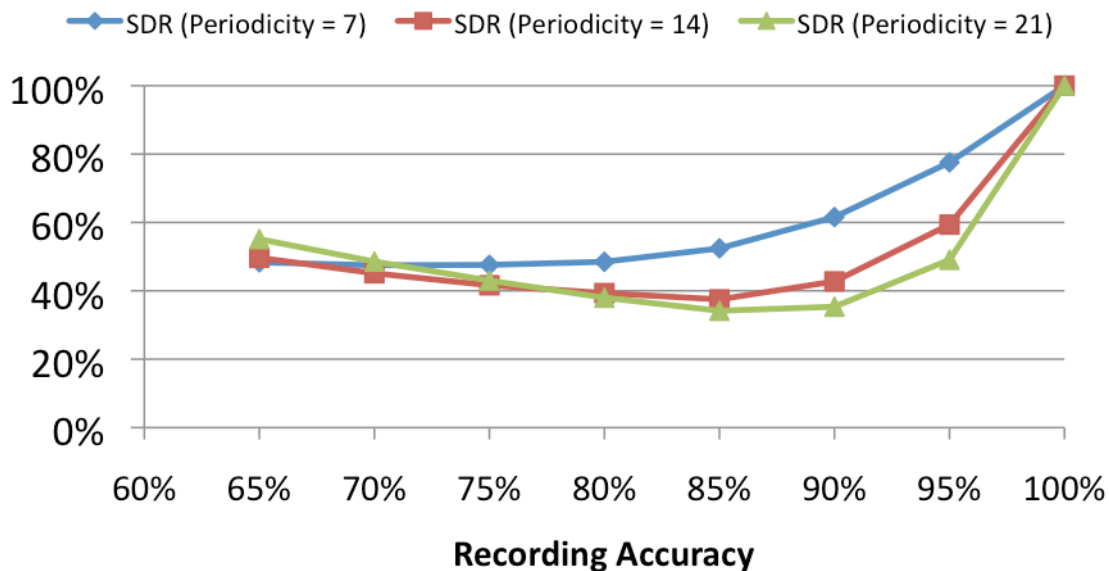


Figure 6-5: Stock-Out Detection Rate as a function of Recording Accuracy, at different Count Periodicities.

To understand this pattern, it is useful to recognize the book stock-out frequency (BSOF) as the product of the physical stock-out frequency (PSOF) and the stock-out detection rate (SDR).

$$BSOF = PSOF * SDR$$

To qualitatively understand how the book stock-out frequency is affected by changes in Recording Accuracy, we focus on the behavior of the stock-out detection rate as Recording Accuracy changes³⁴. Recall that physical stock-outs can be detected and recorded as book stock-outs either:

- (A) when the physical stock-out takes place (i.e., when there are no errors in the inventory record);
- (B) after the physical stock-out takes place, through a periodic count;
- (C), not at all. When (C) occurs, a stock-out takes place, stays undetected and replenishment brings the physical view inventory back to a positive value before the periodic count.

In other words:

$$SDR = 1 - P(C) = P(A) + P(B)$$

If Recording Accuracy is perfect, $P(A) = 1$. In fact, the probability of detecting the stock-out exactly as it occurs, as a function of Recording Accuracy, follows a power law:

$$P(A) = Prob(\text{No Inventory Error})$$

$$P(A) = (\text{Recording Accuracy})^{(\text{Product usage between the last count event and the physical stock-out})}$$

³⁴ The physical stock-out frequency increases as Recording Accuracy decreases, so it is not sufficient to focus on the stock-out detection rate to fully conclude the behavior of the function. In most cases, the relative changes in the stock-out detection rate dominate the changes in the physical stock-out frequency. More importantly, our intent is to present insight into the effect presented rather than a mathematical proof of the simulated results.

When Recording Accuracy is high but not perfect (e.g., 95%), $P(A)$, the probability of detecting a stock-out as it takes place, decreases sharply relative to the perfect recording case. Meanwhile, $P(B)$, the probability of detecting a stock-out during a periodic count is low because stock-out durations are small relative to the periodicity of counts. Therefore, the stock-out detection rate (SDR) decreases sharply relative to the perfect recording case. On the other hand, when Recording Accuracy is low (e.g. 65%), the probability of detecting a stock-out as it occurs (A) is small, but the stock-out duration is much higher relative to the previous case. The stock-out duration is higher because the reorder point is triggered based on the book inventory, which is depleted based on recorded demand, and thus is triggered late. Therefore, $P(B)$, the probability that a stock-out is detected through a periodic count, becomes larger, and so does the stock-out detection rate (SDR). This explains why the stock-out detection rate is not a monotonic function of Recording Accuracy. Because the variations in the stock-out detection rate (SDR) often dominate the change in the physical stock-out frequency (PSOF), the book stock-out frequency (BSOF) is often not a monotonic function of Recording Accuracy.

The fact that the book stock-out frequency is not a monotonic function of Recording Accuracy (or the physical stock-out frequency) implies that it is not a reliable indicator of relative performance over time (i.e., it can provide the wrong answer as to whether this month's service level is better or worse than last month's).

The previous analysis shows that a periodic count policy, while mitigating the effect of imperfect recording by periodically correcting inventory inaccuracies, results in large inaccuracies in both the in-stock probability and stock-out frequency metrics.

6.3.3 Counting on replenishment policy

This section explores the effect of a simple count policy change implemented by the hospital. The change consists of counting the inventory prior to every replenishment transaction, in addition to periodic counts. For instance, inventory is counted every other Monday at 4 pm (14 days

periodicity). As replenishment takes place on Thursday morning, the technician counts the inventory just prior to adding the delivered quantity, and adjusts the book inventory record if a discrepancy is found. If the physical inventory is found to be equal to zero, a book stock-out event is recorded.

The hospital introduced this change to systematize inventory counts, as periodic counts were executed to a varying extent depending on the station and replenishment technicians. Based on the results of interviews, hospital managers were not focused on the problem of undetected stock-out events.

Figure 6-6 compares the scenario of section 6.3.3 with this suggested policy. As expected, because the average count frequency is increased, the number of physical stock-out events diminishes, and the gap between the book and physical stock-out frequencies is virtually eliminated. However, while in steady state no stock-out events are missed, a (variable) delay subsists between the time at which the stock-out event takes place and when it is detected³⁵. For instance, consider the case of a stock-out taking place on July 25th. It is detected as a book view stock-out during the next periodic count on August 4th, and therefore is attributed to the month of August.

³⁵ Because the simulation is only an approximation of the steady state, this delay explains the very small gap between Physical and Book views when always counting on replenishment.

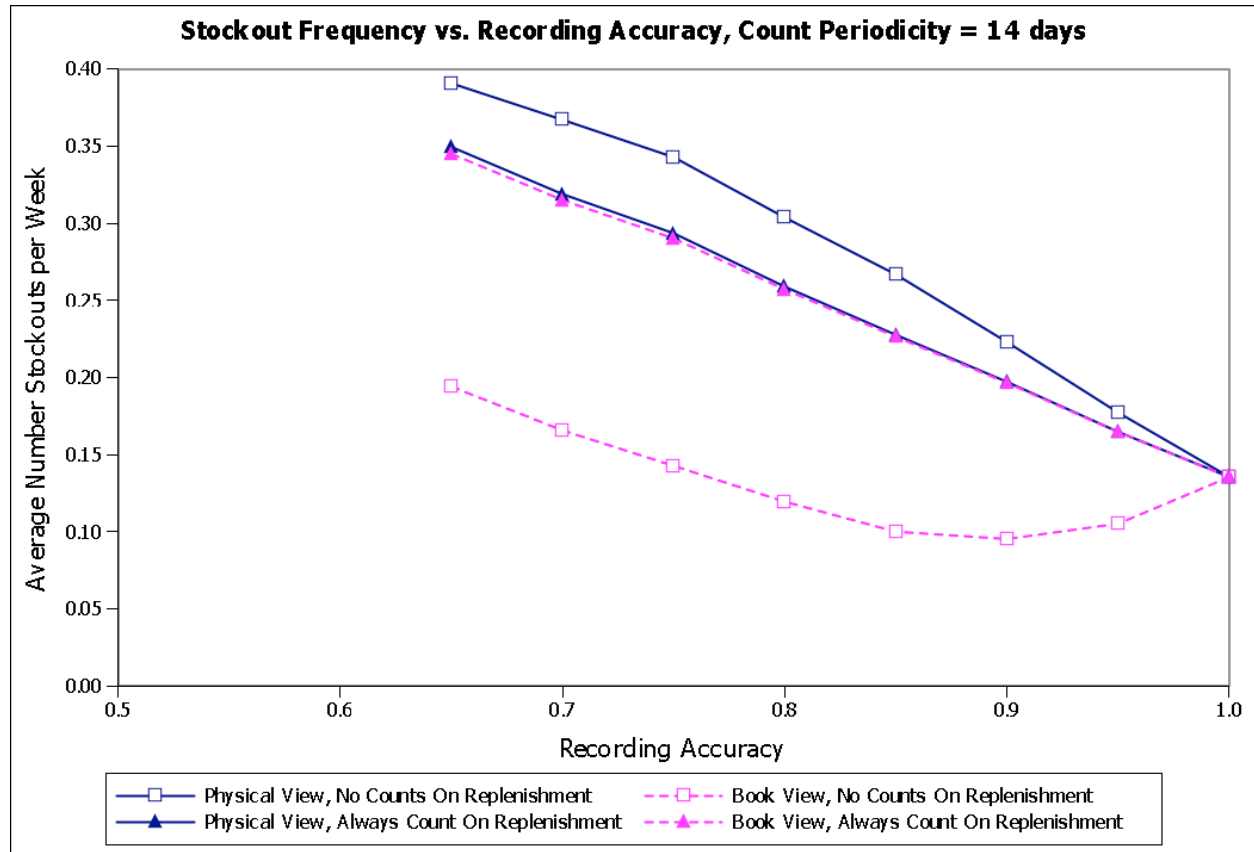


Figure 6-6: Physical and Book Stock-out Frequencies under two scenarios: no counting on replenishment and always counting on replenishment.

Therefore, when calculated every month (or at any periodic interval), the book stock-out frequency metric may reflect stock-out events that occurred the previous month and omit some stock-outs that took place during the period.

6.4 Evaluation of Strategies to Improve Product Availability

A primary advantage of simulation models is that a wide range of policies for inventory inaccuracy mitigation can be tested, both alone and in combination with other policies. To understand

the effects of these policies for a variety of products with different characteristics, a representative panel of products was selected from the set of products used at Lambda.

6.4.1 Scenario definition and performance metrics

Multiple possible scenarios were built around Prevention, Correction and Integration strategies, as summarized in the table below:

Lever	Control Variable	# of Scenarios	Description
Prevention	Recording Accuracy	2	Baseline value Improved accuracy = average(baseline,100%)
Correction	Count On Replenishment (COR)	3	False. Never count on replenishment Partial. Count on 50% of replenishments Full. Always count on replenishment
	Count Timing - Mean - Variability	7	Historical Count Sequence Baseline Gamma Distribution Reduced Variability Gamma Distribution Decreased Mean, Reduced Variability Gamma Distribution Increased Mean Gamma Distribution Increased Mean, Reduced Variability Gamma Distribution
Integration	Dynamic Inventory Policy	2	Baseline: not implemented Active: implemented

Table 6.7: Scenario characteristics. All combinations of the scenarios above were simulated for each item in the panel, corresponding to 84 total inventory inaccuracy mitigation policies.

A full factorial design was employed to explore interaction effects between different scenarios, resulting in 84 alternative configurations. The full results for all products, with all configurations and multiple performance metrics are reported in Appendix I.

Improved Recording Accuracy

The baseline recording accuracy was estimated for each product from ADM transactional data, using the single-item methodology described in section 5.4.2. The improved recording accuracy scenario was defined by dividing the recording accuracy error in half:

$$\text{Improved Recording Accuracy} = 1 - [(1 - \text{Baseline Recording Accuracy})/2]$$

For example, a baseline recording accuracy of 80% would be increased to 90%, and a baseline recording accuracy of 90% would be increased to 95%.

Count on Replenishment

One policy that is sometimes implemented in the hospital is the practice of performing a Count immediately following a Replenishment transaction. Three scenarios are considered: (1) Count is never performed during Replenishment, (2) Count is performed during Replenishment 50% of the time (through random sampling), and (3) Count is always performed during Replenishment. The results indicate that Counting on Replenishment is an inefficient way of improving the Fill Rate compared to simply decreasing the average time between count transactions for the scenarios tested and as shown below.

Count Timing

The time between count transactions (or inter-count time) is estimated from ADM transaction data and modeled using the Gamma distribution, thus allowing for different levels of variability in inter-count times. To validate this modeling choice, the historical count sequence from the ADM transactional data was replayed in the simulation. The system behavior and key service level metrics did not materially differ between these two sets of scenarios, indicating that the Gamma distribution is a good model of the dedicated Count schedule from the transactional data. In all other scenarios, inter-count times are drawn from a Gamma distribution where the mean and coefficient of variation are set equal to those measured from the 15-month transactional data,

after rounding the mean to the nearest nonzero multiple of 3.5 days in order to discretize the search space. Additionally, a “Reduced Variability” scenario is simulated by drawing inter-count times from a Gamma distribution with a coefficient of variation set equal to 0.25. The coefficient of variation of 0.25 is chosen based on the nonlinear effect of inter-count time variability described in section 7.2.2. For this value, count variability has a minimal service level impact and allows for some operational flexibility in executing the count schedule.

Dynamic Inventory Policy

This section examines a simple Dynamic Inventory Policy (DIP) that increases the reorder point according to the calculated value of the average rate of stock loss and the time that has passed since the last confirmed Count event. A confirmed Count event is defined as an instance of a Count on Replenishment when a discrepancy was detected, or as an instance of a scheduled Count event, whether or not a discrepancy was detected. Because it is unknown how often Counts are performed on Replenishments, a Count on Replenishment is assumed not to have occurred unless a discrepancy is detected. At each confirmed Count event, two variables are updated using an exponential moving average (EMA) with smoothing parameter alpha is set equal to 0.10: the average discrepancy size and the average time since between confirmed Count events.

$$\begin{aligned} \text{Average_Discrepancy_Size}(t+1) = \\ \alpha * \text{Observed_Discrepancy}(t+1) + (1 - \alpha) * \text{Average_Discrepancy_Size}(t) \end{aligned}$$

$$\begin{aligned} \text{Average_Time_Between_Conf_Counts}(t+1) = \\ \alpha * \text{Time_Since_Last_Conf_Count}(t+1) + (1 - \alpha) * \text{Average_Time_Between_Conf_Counts}(t) \end{aligned}$$

A discrepancy is defined as the difference between the book inventory level and the physical inventory level. The exponential moving average allows the system to evolve rapidly in time because more recent values are given more weight than values in the past. Then, the average rate of

stock loss is updated by dividing the EMA of the discrepancy size by the EMA of the time between confirmed Count events.

$$\text{Average_Stock_Loss_Rate } (t) = \text{Average_Discrepancy_Size } (t) / \text{Average_Time_Between_Conf_Counts } (t)$$

At each order review period, the expected stock loss is calculated by multiplying the average rate of stock loss by the time that has elapsed since the last confirmed Count event. The book inventory position (i.e. the book inventory level plus any outstanding order quantities) is adjusted by subtracting the expected stock loss from the book inventory level, for the purpose of comparison with the reorder point and the calculation of the order size. The book inventory level is not modified.

In a nutshell, the Dynamic Inventory Policy adjusts the book inventory level at the time of review to account for an expected amount of stock loss, and updates the stock loss rate as information becomes available during count events. It is preferable to increasing the reorder point because it only adds inventory when recording accuracy is low, not all of the time, as an increased reorder point does. This policy can be easily implemented in software given its very limited computational and memory requirements (i.e., it only stores and updates two parameters per Station-SKU).

6.4.2 Sample panel definition

Because 84 alternative configurations are simulated for each Station-SKU, it is desirable to segment Station-SKUs, and to run the simulation on a representative panel of Station-SKUs. This step was performed after conducting the empirical and simulation analyses because they revealed new attributes that had a large effect on the system (e.g. Recording Accuracy). Commonly used techniques such as ABC segmentation focus on unit or dollar volume, which are only some of the attributes that are relevant in the context of the hospital. Based on the hospital materials man-

ager interviews and the insights from the simulation study, the following attributes were considered:

- At the SKU level:
 - o Unit cost,
 - o Commonality: number of stations in which the SKU is stocked,
 - o Storage location: behind a closed cabinet door (CD) or on an open shelf with a radio-frequency button for recording demand (OS)
- At the Station-SKU level:
 - o Mean daily demand,
 - o Coefficient of variation of the demand,
 - o Recording Accuracy,
 - o Mean inter-count time,
 - o Coefficient of variation of the inter-count time

The Partitioning Around Medoids³⁶ (PAM) clustering algorithm was used to construct a panel of representative Station-SKUs using the R statistical package command *pam*. This robust clustering technique minimizes the sum of dissimilarities within clusters and produces clusters defined by each of the representative medoids (Kaufman & Rousseeuw, 1990). In this analysis, the Manhattan distance was used and the mean daily demand was log-transformed to be approximately

³⁶ “Medoids are representative objects of a data set whose average dissimilarity to all the objects in the cluster is minimal. It is similar in concept to means or centroids, but medoids are always members of the dataset.” (Wikipedia, 2010)

normal, and all variables were standardized by calculating their z-score³⁷. The resulting panel is displayed in Table 6.8 below.

Item ID	Station	Description	Storage Type	Mean Daily Demand	CV Daily Demand	Recording Accuracy	Mean IntCt Time	CV IntCt Time
63696	6SE	Sodium Chloride 1000ML .9% Bottle	CD	0.228	2.704	0.957	49.848	0.919
69111	7C	Glove Esteem Nitrile Vinyl of LG	OS	1.504	4.085	0.696	5.738	1.087
64040	6SE	Specimen Container 4OZ	CD	2.322	1.089	0.948	50.094	0.993
10856	OR_SUT	Vicryl Suture 3.0 Box of 24	CD	2.002	1.530	0.798	10.499	0.818
2326	5W	Medivac Tubing	OS	0.020	11.252	0.741	68.825	1.331
56840	LITHO	Stone Tipless Extractor	CD	0.221	3.188	0.969	50.293	1.262
57002	OR_GEN	Cord Monopolar Reusable 10 FT	CD	0.088	8.990	0.826	17.113	1.025
65768	MICU_1	Pack Vas Cath (very bulky)	CD	0.081	3.945	0.876	25.778	2.465
63795	7SE	Bedpan Disposable	OS	0.295	2.503	0.798	22.545	2.039

Table 6.8: Summary of Station-SKU Sample Panel (CD: Closed Door, OS: Open Shelf)

³⁷ i.e., the number of standard deviations above or below the mean of the variable.

6.4.3 Simulation results

For each Station-SKU in the panel, the 84 alternative configurations were simulated with 50 trials³⁸ and a run length of 15 months to equal the length of the ADM data over which the baseline product characteristics were estimated. Because the focus of this section is on actionable recommendations, the real reorder-point and order-up to levels used at Lambda were used, even if they are not optimal given the demand patterns.

The performance of the different strategies was measured using two metrics: the fill rate and the total count frequency (the sum of the number of periodic counts and counts on replenishment per week). The total count frequency is a proxy for the labor cost of the counting process. The plots in Figures 6-6 to 6-9 make it possible to visualize the tradeoff between the service level and the cost for different strategies, and to determine an efficiency frontier.

Based on the simulation results on different products, multiple insights stand out. First, the Reduced Count Variability strategy consistently results in large service level improvements relative to other strategies. Second, when the fill rate is above 99% in the base line because the reorder point is high relative to demand and the recording accuracy is >95%, reducing count variability, improving recording accuracy or implementing the DIP have minimal effects. In this case, the system is offering the required service level to caregivers, and a possible course of action is to optimize costs by considering a moderate reduction in counting frequency and/or the safety stock. Third, when faced with the choice between increasing periodic count frequency or imple-

³⁸ Since the purpose of the simulation was not to construct confidence intervals between different configurations but to assess how the different strategies performed for different types of products, the number of trials was reduced from 200 in the base case to 50 trials, while computing the variance of the estimates to ensure statistical significance.

menting count on replenishment, it is a better use of labor to increase the periodic count frequency. Figures 6-7, 6-8 and 6-9 in the following pages illustrate those results for different products.

Reduced Count Variability

Given that a particular item has sufficient safety stock and good recording accuracy (>80%), the most important corrective measure that can be taken is to reduce the variability in the Inter-count times, i.e. to perform dedicated Counts consistently at some target frequency.

In order to facilitate the visualization of the results, only configurations where the mean inter-count time is equal to the baseline are shown in the plots below, reducing the number of points to $2 \times 3 \times 2 \times 2 = 24$ configurations (or 12 configurations when the reduced count variability is not considered).

For example, item 63795, which corresponds to a disposable bed pan is considered in Figure 6-7. The three vertical clusters of points represent the different Count on Replenishment (COR) levels: as the COR variable increases from 0% to 50% and finally 100%, the total count frequency increases (moving points to the right of the figure) and the fill rate increases (moving points up on the figure). Improving the recording accuracy increases the fill rate by ~ 20 percentage points across the different scenarios. Implementing the DIP further increases the fill rate by another 10 to 20 percentage points across the different scenarios.

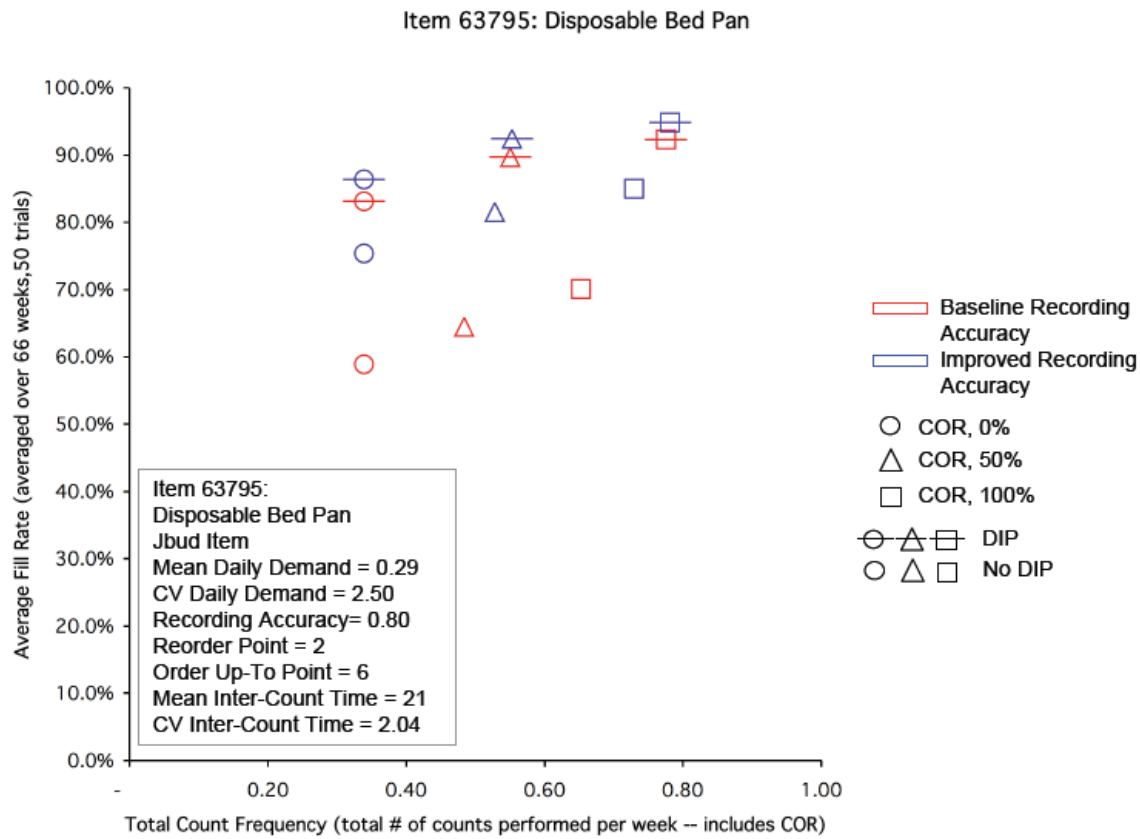


Figure 6-7: Efficiency frontier for item 63795: Disposable Bed Pan.

Figure 6-8 adds the points corresponding to reduced variability to Figure 6-7, and depicts them using full symbols. Reducing the variability of the inter-count times increases the fill rate from the baseline more than increasing the recording accuracy or implementing DIP, and raises it to a fill rate greater than that achieved by both improving accuracy and implementing DIP.

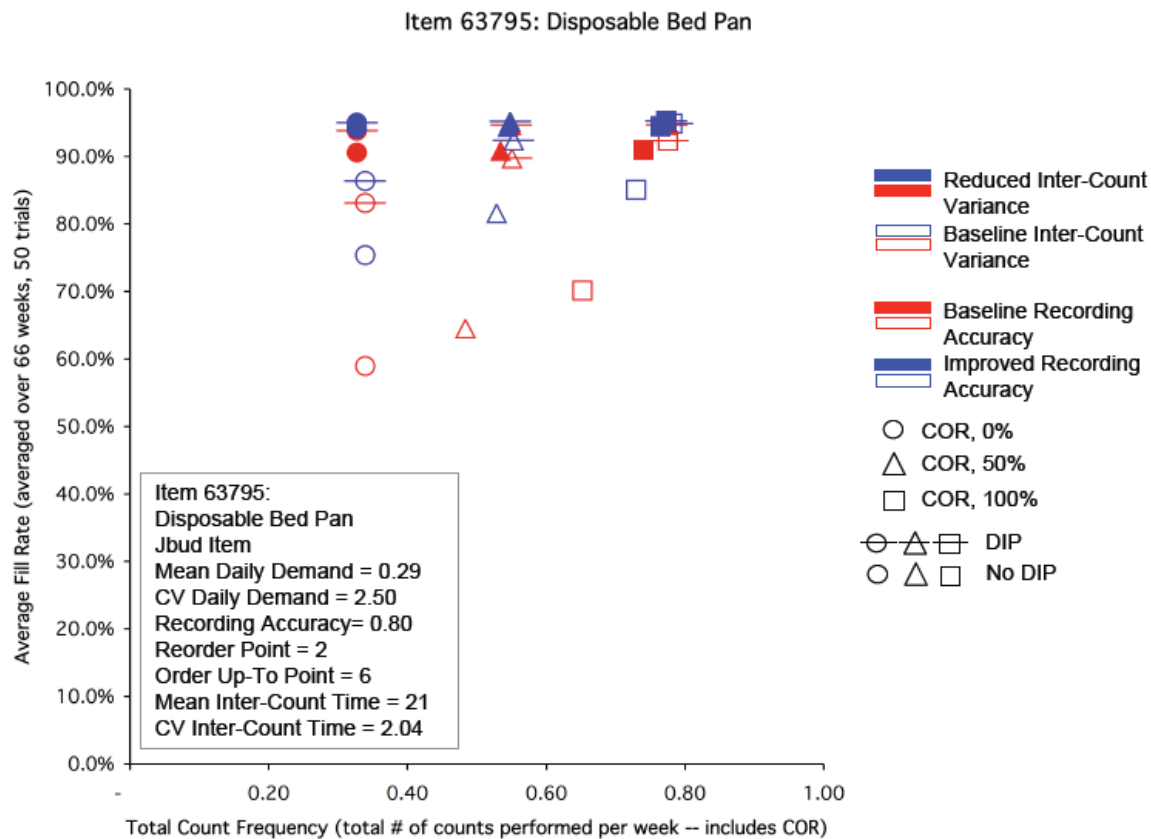


Figure 6-8: Efficiency frontier for item 63795: Disposable Bed Pan. Reduced count variability scenarios are depicted with full symbols.

On the other hand, when the recording accuracy is already very high, the different policies have limited effects. For instance, for item 56840, a tipless stone extractor used in Urology, (Figure 6-9), the fill rate for the base scenario is quite high, at 94% when never counting on replenishment, because the baseline recording accuracy is equal to 97%. Reducing the variability of the inter-count times increases the fill rate to 97% (Figure 6-9). Increasing recording accuracy from 97% to 98.5% may be very difficult to achieve, while decreasing the inter-count variability may be achieved more easily, therefore reducing inter-count variability should be the focus of improvement policies.

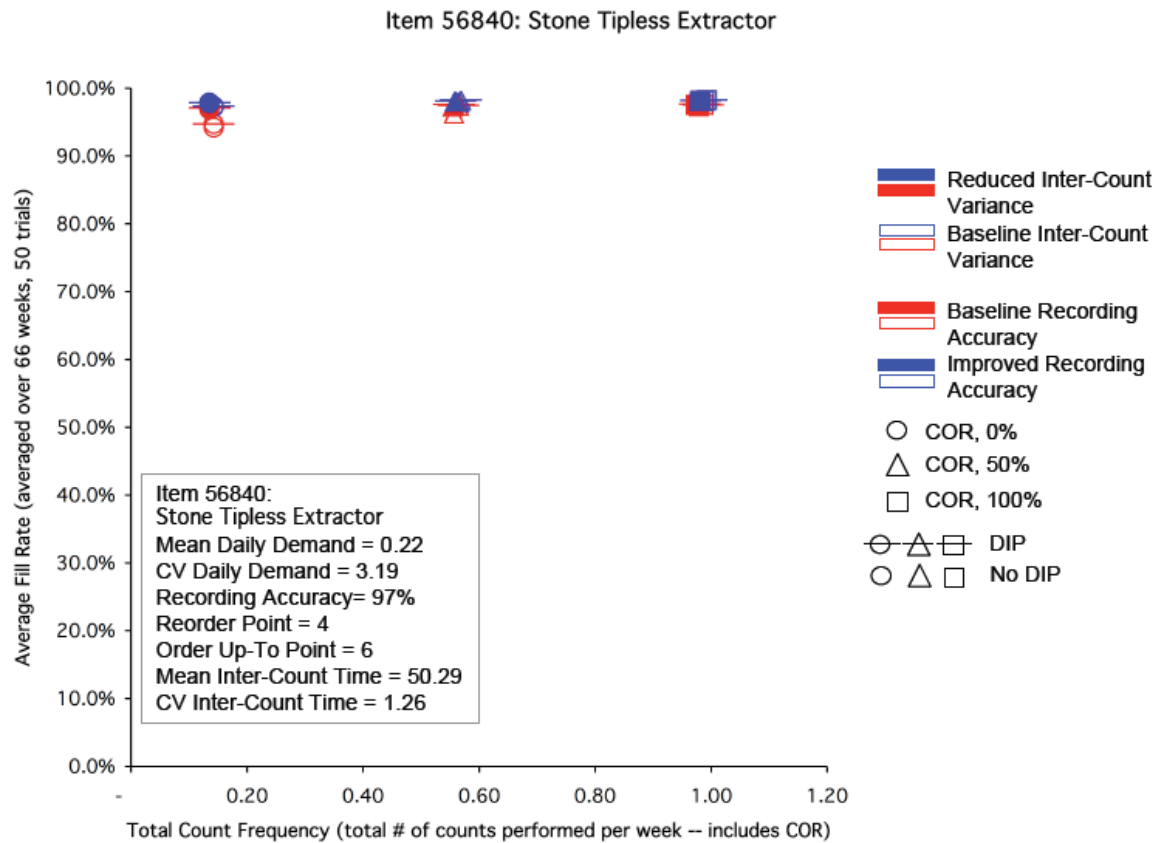


Figure 6-9: Efficiency frontier for item 56840: Stone Tipless Extractor. Reduced count variability scenarios are depicted with full symbols.

Count on Replenishment as An Inefficient Strategy

While Counting on Replenishment removes bias in the (book) stock out frequency metric, its effect on the service level should be compared to the alternative option of increasing the periodic count frequency. Theoretical reasons suggest that Counting on Replenishment is less efficient relative to increasing the periodic count frequency. First, a count performed during replenishment is less likely to detect a discrepancy than a periodic Count because, on average, less time has elapsed since the last count event. Replenishments occur after an order is placed, and in order

for an order to have been placed, the book inventory level must have been at or below the reorder point, which often occurs as the result of a periodic count event. Second, Counting on Replenishment ensures that the book inventory level is accurate when the physical inventory level is at its highest point. Stock outs are likely to occur when the physical inventory is low. Correcting the book inventory level when physical inventory level is high still allows discrepancies to accumulate before the physical inventory level reaches the reorder point.

The simulation model allows us to consider the tradeoff between increasing Count on Replenishment and increasing Count Frequency. To simplify the visualization, the Dynamic Inventory Policy has been omitted from Figure 6-9, although the conclusions are not changed when it is considered. The scenarios where the mean time between counts have been increased or decreased were added to the base line. The figure shows how periodic counts dominate counting on replenishment (i.e. they are further left and further up on the figure), particularly when the variability of the time between counts is reduced (full symbols).

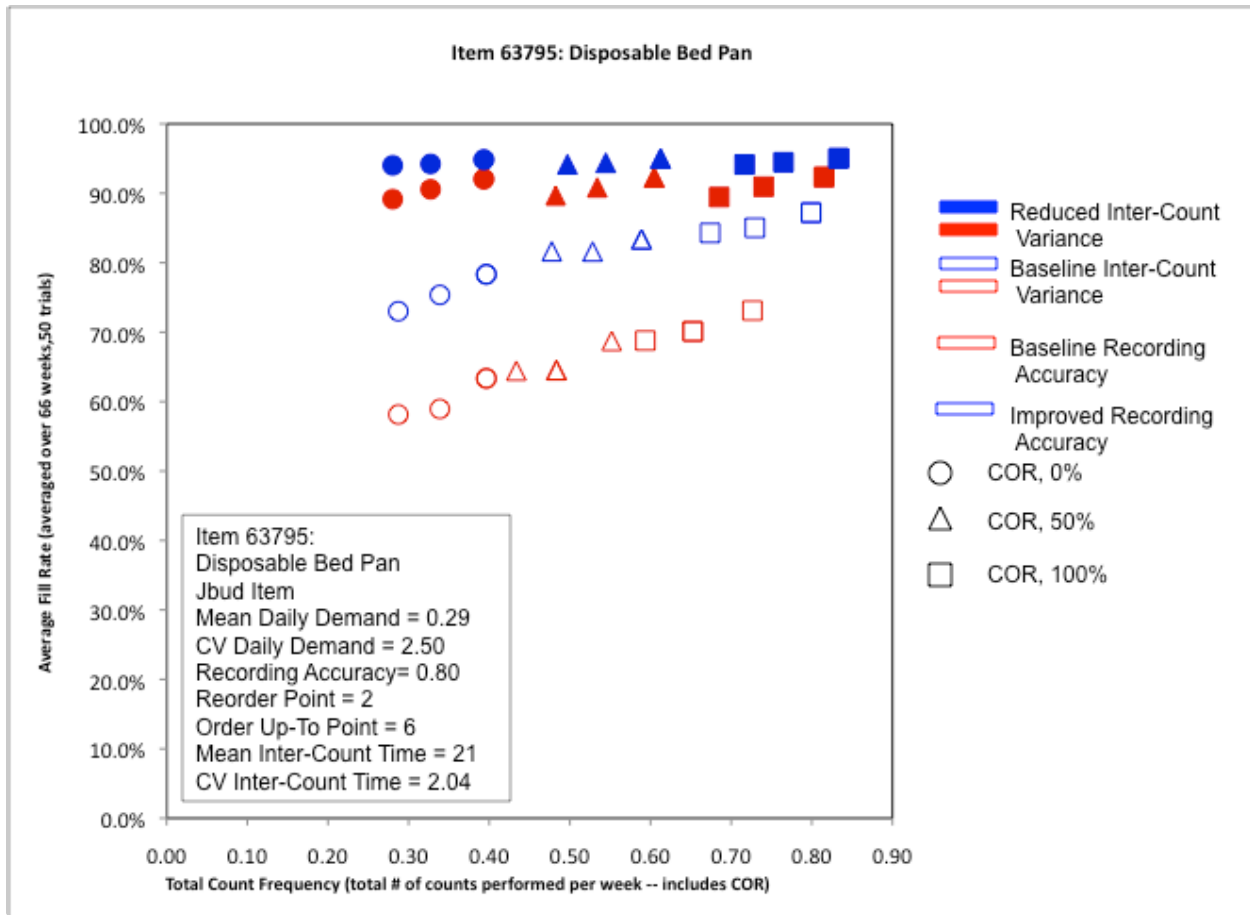


Figure 6-10: Efficiency frontier for item 63795: Disposable Bed Pan. Reduced count variability scenarios are depicted with full symbols.

6.5 Summary

This chapter presents a discrete-event simulation model of the hospital supply chain in the presence of inventory inaccuracies. Exercising the model under a range of Recording Accuracy values and counting assumptions shows that imperfect demand recording causes a large reduction in product availability unless it is mitigated by frequent counts. The tradeoff between the count fre-

quency and the service level is assessed using the model. Variability in the time between counts is identified as a factor that reduces product availability and is investigated in the next chapter.

Second, imperfect demand recording is identified as an important source of bias in performance metrics, causing managers to rely on overly optimistic metrics. Table 6.9 summarizes the situation for different metrics, depending on whether counts occur on replenishment.

Counting on Replenishment	No	Yes
Book view Metric		
Stock-Out Frequency	Biased Doesn't account for Stock-out Duration	Unbiased Doesn't account for Stock-out Duration
In-Stock Probability	Severely Biased	Severely Biased

Table 6.9: Summary of the bias in different product availability metrics.

Chapter 7

Variability in Cycle Counting

In the previous chapter, simulation analysis identified the detrimental effect that variability in cycle counting times has on service levels. This chapter generalizes this insight using an analytical model and quantifies the impact of variability.

To mitigate the detrimental effect of inventory inaccuracies on service levels, managers often resort to cycle counts given their simplicity of implementation. In a standard cycle count program, items are assigned a count frequency based on ABC segmentation. This frequency determines the schedule of counts. After each count, the counted quantity is compared with the book value and if a discrepancy is found, the inventory is recounted and the inventory record is adjusted (Piasecki, 2003). Cycle counts are increasingly being used to address inventory accuracy challenges, with the proportion of retailers intending to perform “more cycle counts” rising from 46% to 78% between 2001 and 2003 (CSA, 2001, 2003). In fact, the proportion of retailers who performed more than ten cycle counts per year increased from 25% in 2004 to 37% in 2005 (Foundation & BearingPoint, 2005). Moreover, DeHoratius & Raman (2008) show using a re-

gression analysis of retail data suggesting that stores which are counted annually have 12% greater inaccuracy³⁹ than stores counted semi-annually.

While managers are correct in believing that increasing count frequencies should in theory correct inaccuracies more promptly and therefore improve service levels, this strategy requires additional resources and increases counting costs. Furthermore, Millet (1994) reports how in a warehousing context, increasing count frequency failed to produce results because operators did not perform counts as diligently as the count frequency increased, and misled management by reporting improved, but ultimately incorrect, accuracy figures. This suggests that behavioral factors play an important role in the successful execution of cycle count programs.

A cycle count program includes a schedule of counts, which may be followed more or less rigorously by materials management staff. This chapter investigates how variability in inter-count times affects product availability and provides empirical data from Lambda documenting such variability.

The outline of this chapter is as follows. Section 7.1 presents the model and demonstrates that if counts are performed independent of the state of the inventory, then for a given target average count frequency, exactly equally spaced counts are optimal in terms of service level. A distribution-free bound is obtained on the optimality gap caused by inter-count time variability in section 7.2.1 and, in section 7.2.3, the performance of this bound as an approximation is evaluated through a numerical simulation. Section 7.3 describes the practical importance of these theoretical findings and shows, using empirical data from a hospital setting, that large variability in inter-count times occurs in practice. Whether this variability could result from dynamic counting

³⁹ After controlling for all other factors at their mean.

policies⁴⁰ is also investigated, as this would violate the assumptions made in section 7.1. Finally, those results are summarized and their managerial implications are discussed in the following chapter (section 8.1.2).

7.1 Service Level Optimality of Equally-Spaced Counts

This section shows that for a fixed number of counts, the optimal timing of counts that maximizes the service level is deterministic and corresponds to equally-spaced counts.

7.1.1 Assumptions and notations

The following five assumptions are made.

A1. The time between successive inventory counts is a positive, i.i.d. random variable T , with finite mean and variance. This assumption requires independence of the state of the inventory (either book or physical), i.e. that no dynamic counting policy is in place⁴⁰.

A2. Counts are perfectly accurate: i.e., immediately after a count the physical and book inventory levels are always equal. Combined with A1, this is equivalent to saying that counts occur according to a renewal process.

⁴⁰ Dynamic counting policies are contingent on the state of the inventory, and can include counting items whose book inventory level is below a threshold or adjusting the inventory record of items found to be out of stock upon visual inspection (Zero Balance Walk).

A3. Inventory inaccuracies have non-negative mean, i.e. on average, the book inventory overstates or is equal to the physical inventory.

This assumption is well justified empirically. Stock loss and theft have been documented in several industries (Kang & Gershwin, 2005) and have the effect of making the book inventory overstate the physical inventory. Furthermore, it seems unlikely that inventory discrepancies, defined as the difference between the book inventory and physical inventory, are on average negative in most practical situations⁴¹. In their empirical analysis of retail inventory inaccuracy, DeHoratius & Raman (2008) find that only one in 37 stores studied had more negative discrepancies than positive discrepancies⁴². The analysis of hospital inventory records showed similar results: negative discrepancies represented on average 7.6% of total discrepancies across 108 stations, with this percentage varying between 0.0% and 28.0%, depending on the station.

A4. The system has reached a steady state. In other words, in evaluating the service level, the average over all initial states is taken (see for instance Hadley & Whitin (1961)). Because of this assumption, the average service level since the last count is defined as $s(t)$, a function of the time since the last count t .

A5. The inventory model is based on the (R, nQ) model described in Hadley & Whitin (1961), where R represents the reorder point and Q the minimum order quantity.

⁴¹ This is theoretically possible, for instance if the only source of inaccuracy is unrecorded product returns.

⁴² From Table C2 of the e-companion provided by these authors, we compute for each store the percentage of records with a positive discrepancy, the percentage of records with a negative discrepancy and calculate the difference (average difference across 37 stores = 11.7%, median = 10.3%, standard deviation = 6.4%). For store 35, 45.8% of records were accurate, 27.0% were positive discrepancies and 27.2% were negative discrepancies, therefore yielding a difference of -0.2%.

This thesis considers the continuous review limiting case. If the book inventory position y_B (book inventory on hand plus on-order minus backorders) is equal to or less than the reorder point R , nQ units are ordered, such that $y_B + (n-1)Q < R + Q \leq y_B + nQ$.

Demand over the lead-time L is i.i.d. Poisson distributed with mean λL and is backordered (i.e. there are no lost sales). Inventory inaccuracies per unit time are i.i.d. and are modeled as the difference of two independent Poisson random variables, which results in a Skellam distribution (Skellam, 1946). This discrete distribution was used in previous studies of inventory inaccuracy (DeHoratius, et al., 2008), and has the advantage of allowing for both positive and negative inaccuracies and being easily fitted through the method of moments. Furthermore, recent extensions show that the Skellam distribution can also be interpreted as the difference between two correlated Poisson variables, allowing for the modeling of correlation between positive and negative inaccuracies through the use of modified parameters (Karlis & Ntzoufras, 2006).

The following notations are used:

Positive scalars:

t – Time since last count

R – Reorder point

Q – Minimum order quantity

L – Deterministic lead time

λ – Demand rate

ε_1 – Rate of unrecorded additions of stock

ε_2 – Rate of unrecorded reductions of stock

$$\delta = \frac{\varepsilon_1}{\varepsilon_2}$$

Random variables:

$X_1(t)$ – Accumulated positive inventory inaccuracies since the last count

$X_2(t)$ – Accumulated negative inventory inaccuracies since the last count

D_L – Demand over the lead time L

Y_B – Recorded inventory position (recorded inventory on hand plus on order minus backorders)

Z_B – Recorded inventory on hand

Z_p – Actual inventory on hand (recorded inventory on hand plus accumulated inaccuracies since the last count)

T – Time between successive count events

7.1.2 Derivation

The derivation is structured in two parts: first, for the inventory model assumed in A5, the in-stock probability is shown to be a decreasing function of the time since the last count t ; this result and Jensen's inequality are then used to show that equally spaced counts maximize the service level.

Result 1: If $R > \lambda L \delta$, then $p(t)$, the probability of having positive physical on-hand inventory as a function of the time since the last count t , is a decreasing function for the assumed inventory model.

Observation 1: Service level, defined as the percentage of time in stock, can be expressed as a function of $p(t)$:

$$S(T) = \frac{1}{E[T]} E_T \left[\int_{t=0}^{t=T} p(t) dt \right]$$

For the Poisson demand model assumed in (A5), Poisson arrivals see time averages and therefore $S(T)$ also corresponds to the demand fill rate.

When safety stock is positive (i.e., $R > \lambda L$), the inequality $R > \lambda L \delta$ is satisfied. It is therefore not a restricting condition because inventory systems generally require positive safety stocks in order to ensure a high service level.

Based on the results of the extensive numerical simulations performed, it is conjectured that if A1–A4 hold, then result 1 holds for a variety of other ordering policies, including in the presence of lost sales, as long as inventory inaccuracies have a nonpositive mean⁴³.

Derivation of Result 1

For any time t , Hadley & Whitin (1961) show that the inventory position is uniformly distributed over $[R + I, R + Q]$ ⁴⁴. Because all decisions are based on the book inventory position, under the assumption that inventory inaccuracies are not affected by stock-outs, this result is still valid.

⁴³ Intuitively, if inventory inaccuracies have a positive mean, then the service level increases as the time since the last count increases, causing excess inventory.

⁴⁴ The authors derive this result for the inventory position immediately after the review, but here we consider the limiting case of continuous review.

Physical inventory on hand is equal to book inventory on hand plus the accumulated inventory inaccuracies since the last count, as follows.

$$Z_P = Z_B + X_1 - X_2 = Y_B - D_L + X_1(t) - X_2(t)$$

$$Y_B \sim Unif[R+1, R+Q]$$

$$D_L \sim Poisson[\lambda L]$$

$$X_1(t) \sim Poisson[t\varepsilon_1]$$

$$X_2(t) \sim Poisson[t\varepsilon_2]$$

D_L follows a Poisson distribution with rate λL . Positive (X_1) and negative (X_2) inventory inaccuracies are assumed to be independent of each other and of the demand over the lead time D_L . Accumulated inventory inaccuracies are therefore distributed according to a Skellam distribution, i.e., the difference between a Poisson random variable of rate $t\varepsilon_1$ and an independent Poisson random variable with rate $t\varepsilon_2$. Because the sum of two independent Skellam distributions is a Skellam distribution with parameters respectively equal to the sums of the original parameters, $N(t) = D_L - X_1(t) + X_2(t)$ is Skellam distributed with parameters $(\lambda L + t\varepsilon_2, t\varepsilon_1)$ and c.d.f. $F_{N=Skellam(\lambda L+t\varepsilon_2, t\varepsilon_1)}$. Because the timing of the inventory counts is independent of inventory levels, the in-stock probability $p(t)$ is equal to:

$$\begin{aligned} p(t) &= P[Z_P(t) > 0] = P[Y_B - N(t) > 0 | t] = P[N(t) < Y_B | t] \\ p(t) &= \sum_{k=1}^Q P[Y_B = k] P[N(t) < R + k | Y_B = k] \\ p(t) &= \sum_{k=1}^Q \frac{1}{Q} P[N(t) < R + k | Y_B = k] = \frac{1}{Q} \sum_{k=1}^Q F_{N(t)=Skellam(\lambda L+t\varepsilon_2, t\varepsilon_1)}(R+k-1) \end{aligned}$$

From the expression above, when inventory inaccuracies can only be negative (i.e., $\varepsilon_1 = 0$ and $\varepsilon_2 > 0$), $N(t)$ is a Poisson distribution with rate $\lambda L + t\varepsilon_2$. The Poisson distribution is strictly stochastically ordered using the usual stochastic order, as defined below.

$$\begin{aligned}
 N(t_1) &\sim \text{Poisson}[\lambda L + t_1 \varepsilon_2] \\
 N(t_2) &\sim \text{Poisson}[\lambda L + t_2 \varepsilon_2] \\
 t_2 &> t_1 \\
 \Leftrightarrow \forall x \in [-\infty, +\infty], F_{N(t_2)}(x) &< F_{N(t_1)}(x) \\
 \Leftrightarrow N(t_1) &\prec N(t_2)
 \end{aligned}$$

In this simplified case, Result 1 follows from this sufficient condition.

However, assumption A3 is more general in that inventory inaccuracies are nonpositive on average. The case in which positive inventory inaccuracies as well as negative inventory inaccuracies occur is now considered, i.e., $0 < \varepsilon_1 \leq \varepsilon_2$. Unfortunately, first-order stochastic ordering is not conserved in this case⁴⁵, thus the proof of Result 1 in this general case relies on careful analysis of the cumulative distribution function and probability density function of the Skellam distribution, which is provided in Appendix E.

Result 2: For fixed $\mu = E[T]$, the service level is maximized for this inventory model when cycle counts are deterministic and equally spaced ($\sigma = \sqrt{\text{Var}[T]} = 0$).

Intuitively, the service optimality of equally spaced counts can be understood using a single-period model, with a period length of τ time units. A budget of two counts for the entire period is set, and the first count must occur at the beginning of the period (therefore, the period starts with no inventory inaccuracies). The timing of the second count divides the time interval into two

⁴⁵ It was verified numerically that for instance, when $\lambda L = 1$, $\varepsilon_1 = 0.05$, $\varepsilon_2 = 0.10$, $t_2 = 10$, $t_1 = 1$, $N(t_1)$ is not stochastically greater than $N(t_2)$.

subperiods. If they are of unequal length, service levels suffer in two ways: first, one subperiod is longer and therefore inventory inaccuracies are more likely to have a detrimental effect; second, the likelihood of correcting an error at the end of the shorter subperiod is reduced (Figure 7-1).



Figure 7-1: Simplified timeline of count events

Since the time of the second count t can vary, it can be set to maximize the average service level s .

$$s(t) = \frac{1}{t} \int_0^t p(u) du, t > 0$$

The mean inter-count time μ is fixed as it does not depend on t , but the variance of inter-count times σ^2 is a function of t :

$$\mu = \frac{1}{2} [(t) + (\tau - t)] = \frac{\tau}{2}$$

$$\sigma^2 = \frac{1}{2} [(t)^2 + (\tau - t)^2] - \mu^2 = \frac{1}{4} \left(t - \frac{\tau}{2} \right)^2$$

The average service level during the period is now considered.

$$S(t) = \frac{1}{\tau} [ts(t) + (\tau - t)s(\tau - t)] = \frac{1}{\tau} \left[\int_0^t p(u) du + \int_0^{\tau-t} p(u) du \right] = \frac{1}{\tau} [r(t) + r(\tau - t)]$$

$$S'(t) = \frac{1}{\tau} [p(t) - p(\tau - t)]$$

$$S'(t) = \begin{cases} 0 & \text{when } p(t) = p(\tau - t), \text{ i.e. when } t = \frac{\tau}{2} \\ > 0 & \text{when } t < \frac{\tau}{2} \\ < 0 & \text{when } t > \frac{\tau}{2} \end{cases}$$

Because the service level p is a strictly decreasing function of t , $t^* = \tau / 2$ maximizes the average expected in-stock probability over the period.

Derivation of Result 2

For notation convenience, let $\mu = E[T]$ be the average inter-count time.

Let $r : t \rightarrow r(t) = \int_0^t p(u) du$. By noting that $r'' = p'$ and that p is decreasing (Result 1), it follows that r is concave and therefore Jensen's inequality applies:

$$r(E[T]) \geq E[r(T)]$$

$$r(\mu) \geq E[r(T)]$$

$$S(\{\mu\}) \geq \frac{1}{\mu} E[r(T)] = S(T)$$

This concludes the proof of the optimality by showing that the service level achieved with equally spaced counts (corresponding to the left term in the inequality above) is always larger than the service level $S(T)$. The optimum is unique when p is strictly decreasing because of the strict concavity of r . Beyond Result 1, the proof relies on assumptions A1–A4 about the nature of

the inventory inaccuracies and the count process, but not directly on the specifics of the inventory model used.

7.2 Quantifying the Service Level Deterioration Resulting From Variability in Inter-count Times

This section seeks to quantify analytically the service level deterioration as a result of the variability in inter-count times. The approach seeks a bound on $S(T)$ as a function of the mean and variance of T using convexity. The tightness of the bound obtained is then investigated using a numerical analysis.

7.2.1 Derivation

From Jensen's inequality, Agnew (1972) shows the following theorem:

“Let X be a nonnegative random variable with $E(X) = \mu > 0$

and $E(X^2) = \lambda = \mu^2 + s^2 < +\infty$.

Suppose that $f: [0, +\infty) \rightarrow \mathbf{R}$ with $f(0) = 0$ and $g(x) = f(x)/x$ convex on $(0, +\infty)$.

Then, $E(f(X)) \geq \mu g(\lambda/\mu) = (\mu^2/l) f(\lambda/\mu)$ and the bound is sharp.”⁴⁶

⁴⁶ Note that in this citation of Agnew (1972), λ is *only* a notation for $E(X^2)$ and is *not* related to the demand rate previously modeled.

If $X = T$, $f = r$ (respectively $-r$) is chosen, it can be verified that f satisfies the hypotheses of the above theorem iff s is convex (respectively concave). If $s: t \rightarrow s(t) = \frac{1}{t} \int_0^t p(u) du, t > 0$ is convex,

then:

$$\mu S(T) = E[Ts(T)] \geq \mu s\left(\frac{\mu^2 + \sigma^2}{\mu}\right) = \mu s(\mu_{eq})$$

$$S(T) \geq s\left(\frac{\mu^2 + \sigma^2}{\mu}\right) = s(\mu_{eq})$$

$$\mu_{eq} = \mu \left(1 + \left(\frac{\sigma}{\mu}\right)^2\right) = \mu(1 + CV^2)$$

with the inequality reversed if s is concave.

Renewal theory suggests that this bound is also an approximation of $S(T)$. Let m be the mean *residual* time between the instant an inaccuracy is introduced to the instant it is corrected through a count; μ is the mean and CV is the coefficient of variation of the inter-count times. Renewal theory of “recurrence times” states (Cox, 1962):

$$m = \frac{\mu}{2}(1 + CV^2) = \frac{\mu_{eq}}{2}$$

The service level can be approximated by $s(\mu_{eq})$ under the additional assumption that the service level only depends on the mean *residual* time of an inventory inaccuracy m , as opposed to the full distribution of inter-count times.

The lower (upper) bound result depends on s being convex (concave). To determine whether the service level is above or below the bound $s(\mu_{eq})$, the convexity of the average in-stock probability s as a function of the time since the last count event needs to be examined. For (R, Q) inventory policies and in the absence of inventory inaccuracies, Zipkin (1986) shows that average out-

standing backorders are jointly convex in order size, reorder point and standard deviation of lead time demand. To the best of this thesis' author knowledge, the literature does not consider the convexity of service level metrics under inventory inaccuracy.

Convexity of s means that the average service level deteriorates slower as the time since the last count t increases, while concavity implies faster deterioration as t increases. Therefore, if average service level has *diminishing* marginal deterioration in t (s is *convex* over the range spanned by the distribution of inter-count times), then the average service level is *superior to* the service level obtained for equally spaced counts for μ_{eq} , which is the equivalent count periodicity in the absence of variability. It is reasonable to expect that s reaches a limit as the time since the last count approaches infinity, and since it is a decreasing function, it is likely, although not mathematically necessary, that it is convex over a domain $[t_{critical}, +\infty)$.

In Appendix F, this result is demonstrated analytically for a well-cited closed-form model of service level under inventory inaccuracies: the Minimum Actual Protection Level (MAPL) developed by Morey (1985). A simple closed-form expression for $t_{critical}$ is obtained, which depends only on two parameters that have a practical interpretation: the target Type I service level in the absence of inventory inaccuracies, α , and the ratio of the variance of lead time demand to the variance of inventory inaccuracies, $\gamma = \frac{\sigma_{DL}^2}{\sigma_{\epsilon}^2}$. Based on this result, it is conjectured that, in most operational settings, $s(\mu_{eq})$ constitutes a lower bound to the actual service level. This conjecture is subsequently verified through a numerical analysis, which is reported in Appendix G.

7.2.2 Count Inefficiency metric

The lower bound $s(\mu_{eq})$ only uses a modified inter-count time to account for the effect of variability. This means that the Count Inefficiency, defined as the percentage of counts that could have been avoided if all counts were equally spaced, can be expressed as a function of the square of the coefficient of variation alone, and follows an S-curve shape:

$$\mu_{eq} = \mu \left(1 + \left(\frac{\sigma}{\mu} \right)^2 \right) = \mu (1 + CV^2)$$

$$\text{Count Inefficiency} = 1 - \frac{\text{Number of Counts Necessary per unit time}}{\text{Number of Counts Performed per unit time}}$$

$$\text{Count Inefficiency} = 1 - \frac{1/\mu_{eq}}{1/\mu} = \frac{\mu_{eq} - \mu}{\mu_{eq}} = \frac{CV^2}{1 + CV^2}$$

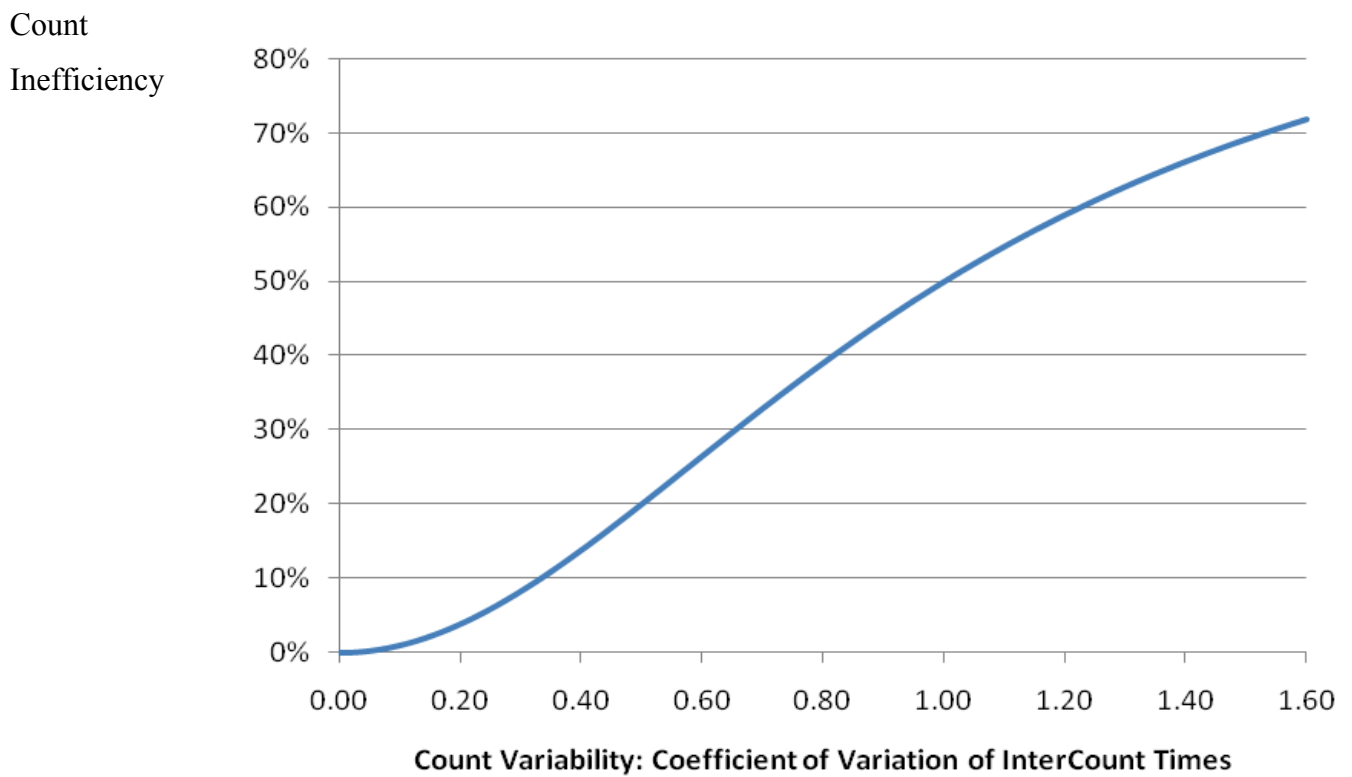


Figure 7-2: Count Inefficiency as a function of Count Variability, showing an S-curve pattern

The Count Inefficiency formula provides managers with a rapid estimate of how count variability reduces their efficiency in the sense that the same result could have been achieved with fewer equally spaced counts. The relationship between the coefficient of variation of inter-count times

and Count Inefficiency is nonlinear: Count Inefficiency increases slowly for $CV < 0.2$, reaches 20% for $CV = 0.5$, and is as high as 50% for $CV = 1.0$ (see Figure 7-2).

7.2.3 Numerical study

This section attempts to quantify the service level impact of inter-count time variability and separates the effect of inventory inaccuracy and imperfect timing of cycle counts. Morey's Minimum Average Protection Level (MAPL)⁴⁷ is used as a service level metric (i.e., $s = s_{MAPL}$), and the expected service level $s_{MAPL}(T)$ is compared with its bound $s(\mu_{eq})$ for different scenarios using Monte Carlo simulation. The scenarios reflected a wide range of operational parameters spanning those found in the hospital data.

The parameters are normalized to simplify the analysis. Let:

- α : target MAPL (i.e., Type I service level) in the absence of inventory inaccuracies
- μ : mean time between counts
- CV : coefficient of variation of the time between counts
- $\mu_{eq} = \mu \left(1 + \left(\frac{\sigma}{\mu} \right)^2 \right) = \mu (1 + CV^2)$
- γ : ratio of the variance of the inventory discrepancy σ_ϵ^2 to the variance of the lead time demand σ_{DL}^2

⁴⁷ The MAPL is defined as the probability that the sum of the lead time demand and maximal inventory inaccuracy during the replenishment cycle does not exceed the reorder point. Therefore, it is a conservative estimate of the Type I Service level. It is described formally in Appendix F.

Therefore, by adapting the expression given in Morey (1985), key expressions can be rewritten as follows.

$$MAPL(\alpha, \gamma, t) = 2 \cdot \Phi\left(\frac{\Phi^{-1}(\alpha)}{\sqrt{1 + t\gamma}}\right) - \alpha$$

$$t_{critical}(\alpha, \gamma) = \frac{1}{\gamma} \left(\frac{1}{3} (\Phi^{-1}(\alpha))^2 - 1 \right)$$

Using common random numbers for each configuration (Law, 2007), 10,000 inter-count times are sampled from a Gamma distribution, which is a common two-parameter model of interarrival times. Figure 7-3 plots the simulated service level as a function of the mean time between counts for two different values of the coefficient of variation (CV).

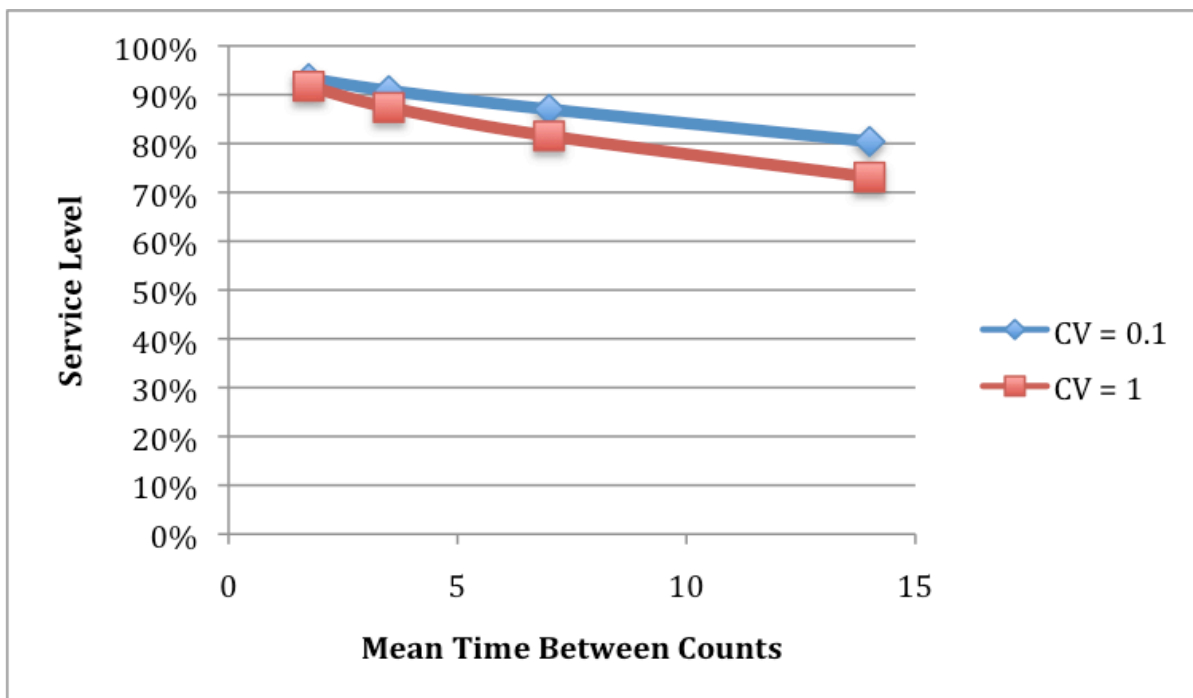


Figure 7-3: Service Level as a function of the Mean Time between Counts

Then, a comparison is made among:

- α : target MAPL in the absence of inventory inaccuracies
- $S_{MAPL}(\mu)$: MAPL that would have been achieved *without* variation in the inter-count times
- $S_{MAPL}(T)$: MAPL achieved with variation in the inter-count times, obtained through numerical simulation
- $S_{MAPL}(\mu_{eq})$: MAPL bound derived in section 7.2.1

To simplify the presentation of results and reduce the number of parameters to investigate, $K = \gamma \mu$ is held constant to maintain a constant service level assuming equally spaced counts for a given target service level α . This reflects managerial action, which aims to maintain an acceptable service level by decreasing the average time between counts (m) as inventory inaccuracy (γ) increases. Results for a value of $K = 0.5$, which is equivalent to counting exactly every 50 days when the variance of asset errors represents 1% of the variance of lead time demand⁴⁸, are presented. Sensitivity analyses show that as K increases, the service level losses as a result of inventory inaccuracy and count variability increase, as well as the gap between the service level achieved and the lower bound⁴⁹.

⁴⁸ Also, Morey (1985) presents a numerical example for which $K = 0.506$, and the service level deteriorates from 94.6% to 86.1% given inventory inaccuracy.

⁴⁹ Results for different values of K are presented in Appendix G.

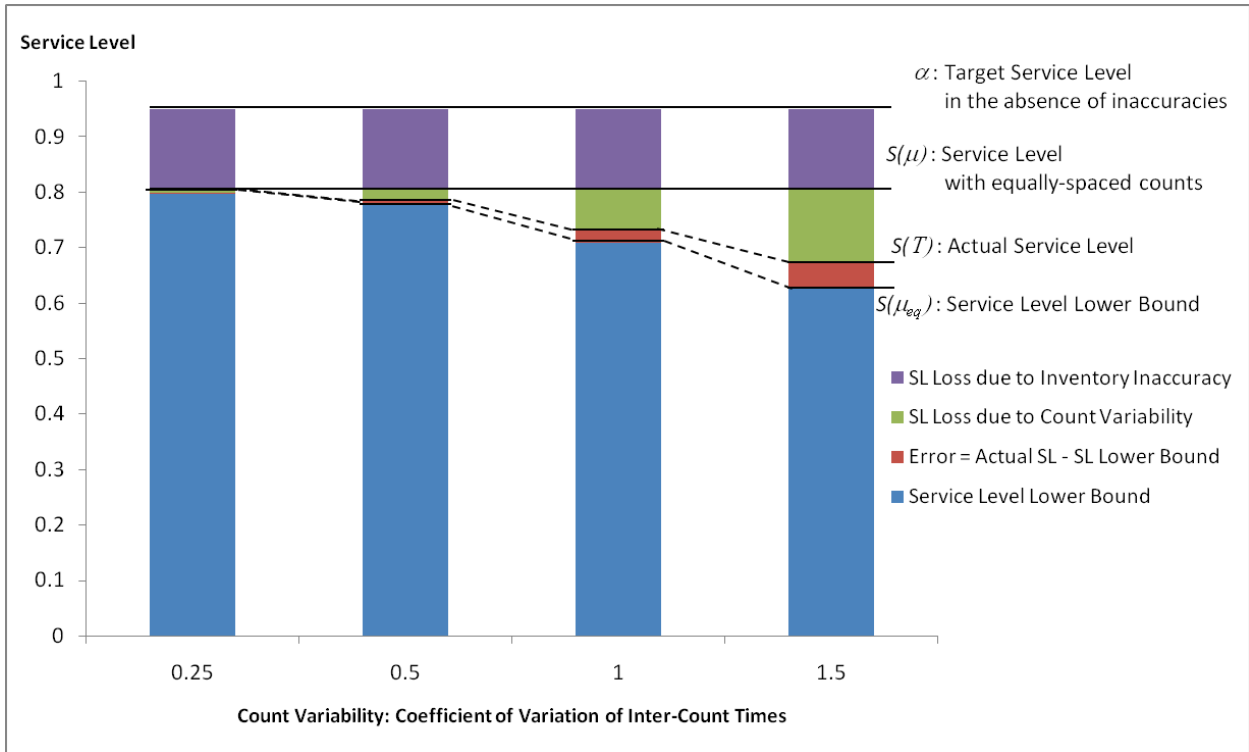


Figure 7-4: Breakdown of Service Level into its different values as a function of Count Variability ($K = 0.5$, $\alpha = 0.95$). As inter-count time variability increases, the service level impact of inequally spaced counts increases and the bound is less tight.

The tightness of the bound $S_{MAPL}(\mu_{eq})$ is investigated and is found to provide a close approximation of $S_{MAPL}(T)$ as long as the coefficient of variation of inter-count times is below one, as seen in Table 7.1. When the coefficient of variability of inter-count times is around one, the service level degradation resulting from count variability is of the same order as the loss resulting from inventory inaccuracy (Table 7.1), showing that inter-count time variability can undermine counts as a mitigation mechanism.

α	CV	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.80					
	0.10	0.7073	0.0001	0.0006	0.0920
	0.25	0.7036	0.0010	0.0034	0.0920
	0.50	0.6909	0.0043	0.0129	0.0920
	1.00	0.6482	0.0147	0.0451	0.0920
	1.50	0.5966	0.0276	0.0838	0.0920
0.90					
	0.10	0.8038	0.0001	0.0007	0.0954
	0.25	0.7996	0.0009	0.0041	0.0954
	0.50	0.7853	0.0040	0.0154	0.0954
	1.00	0.7352	0.0146	0.0549	0.0954
	1.50	0.6711	0.0300	0.1035	0.0954
0.95					
	0.10	0.8700	0.0001	0.0007	0.0793
	0.25	0.8662	0.0006	0.0039	0.0793
	0.50	0.8531	0.0027	0.0150	0.0793
	1.00	0.8052	0.0110	0.0545	0.0793
	1.50	0.7400	0.0257	0.1050	0.0793
0.98					
	0.10	0.9259	0.0000	0.0005	0.0536
	0.25	0.9230	0.0002	0.0032	0.0536
	0.50	0.9128	0.0010	0.0125	0.0536
	1.00	0.8736	0.0055	0.0474	0.0536
	1.50	0.8151	0.0170	0.0944	0.0536
0.99					
	0.10	0.9521	0.0000	0.0004	0.0375
	0.25	0.9499	0.0000	0.0026	0.0375
	0.50	0.9420	0.0002	0.0103	0.0375
	1.00	0.9100	0.0021	0.0404	0.0375
	1.50	0.8590	0.0106	0.0829	0.0375

Table 7.1: Service Level Breakdown for different values of α , the target service level in the absence of inventory inaccuracies ($K = 0.5$).

7.3 Empirical Study of Inter-count Times

The theoretical work developed in the prior sections suggests that reducing the variability of inter-count times is a significant opportunity for the efficiency of a cycle count program. This section investigates variability in inter-count times using Lambda's data in subsection 7.3.1, and studies it in more detail by focusing on the Cardiac Catheterization lab in subsection 7.3.2.

However, the possibility exists that this variability results not from the imprecise execution of a cycle count program but rather from dynamic counting policies, which violate the assumptions of the analytical model presented in section 7.1 and can potentially improve performance (DeHoratius, et al., 2008; Kok & Shang, 2007). Ideally, one would like to quantify empirically the effect of inter-count time variability on service levels. Unfortunately, no reliable data on service levels were available to us because true inventory levels are only known at the time of counts; therefore, inventory inaccuracies themselves introduce bias in service level metrics calculated from recorded data (see previous chapter). As an alternate approach, a statistical procedure to test for the presence of a class of dynamic counting policies is proposed and implemented in subsection 7.3.3

Finally, the potential for reducing the overall number of counts while maintaining current service levels is estimated in subsection 7.3.4.

7.3.1 Measuring Inter-count Time Variability

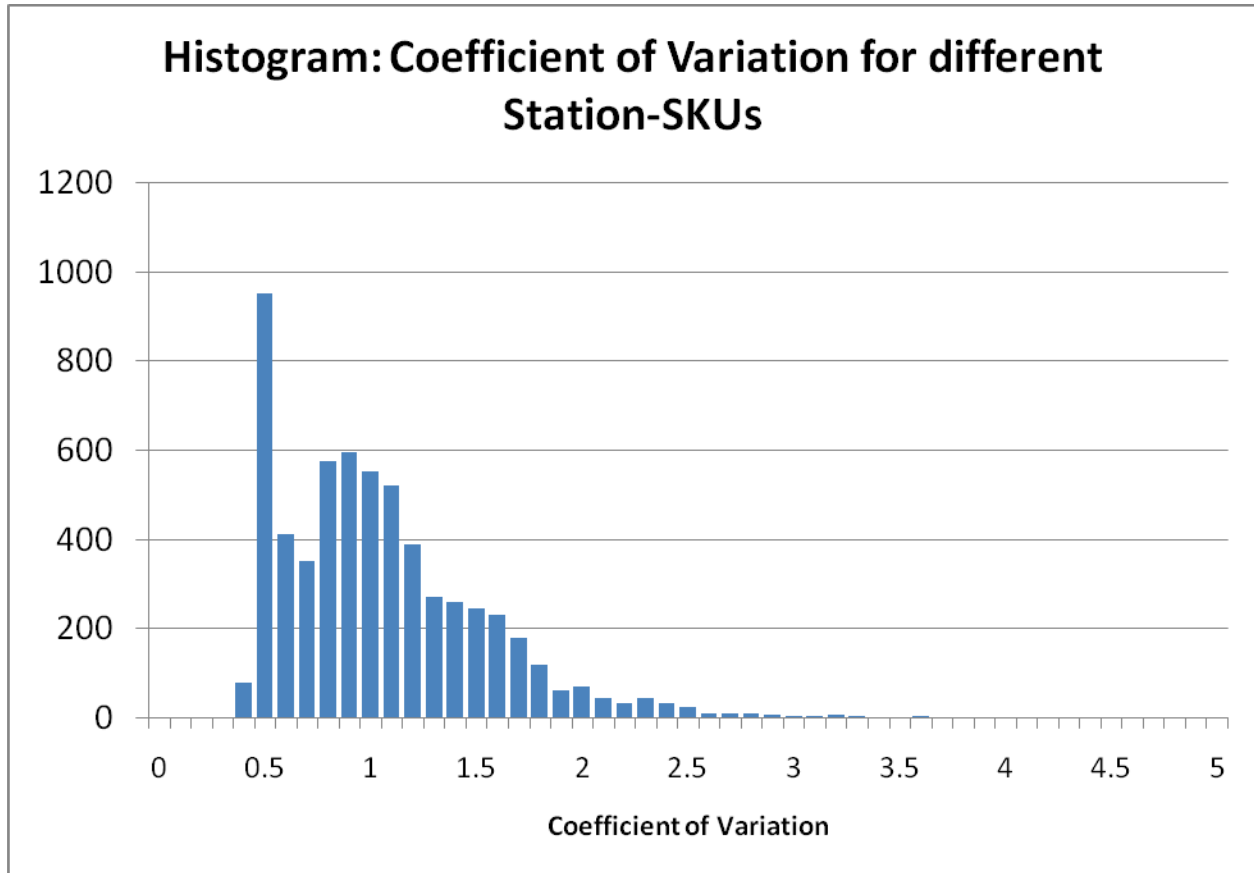
For each Station-SKU, the variability of inter-count times is characterized through its coefficient of variation. The coefficient of variation of inter-count times is not defined if less than three counts were performed during the study period (i.e. two time differences are available), and is not meaningful when it is calculated with less than 10 data points.

Table 7.2 shows that a moderate increase in variability as the number of counts performed during the study period increases. Regardless of the cut-off criteria used, the average coefficient of variation was always above 0.8, which implies large values of the count inefficiency metric.

Number of counts per Station-SKU	Number of Station SKUs	Total number of counts	Average C.V.
0	1,816	0	N/A
1	3,035	3,035	N/A
2	2,950	5,900	N/A
3	1,849	5,547	0.83
4	1,512	6,048	0.92
5	1,039	5,195	1.00
6	935	5,610	1.02
7	798	5,586	1.02
8	618	4,944	1.06
9	510	4,590	1.07
10	420	4,200	1.10
>10	6,110	261,934	1.10
Total	21,592	312,589	

Table 7.2: Average Coefficient of Variation of inter-count times broken down by the number of cycle counts performed during the 15-month study period

For the purpose of ensuring that a sufficient number of data points are used in estimating variability, Station-SKUs counted at least ten times during the 15-month study period (6,110 out of 21,587) were selected. In this sample, a high level of variability of inter-count times was found, as shown by the histogram plot of the coefficient of variation in Figure 7-5. The distribution of the coefficient of variation across different Station-SKUs is bimodal, with a distinct first mode at CV = 0.5 and another mode at CV = 0.9. The first mode contains 954 Station-SKUs, of which 922 belong to the three Cardiac Catheterization Lab stations that stock a total of 1,645 Station-SKUs.



Coefficient of Variation	<0.4	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	>2.1
# of Station SKUs	2	82	954	412	352	577	595	553	520	388	272	261	244	230	180	118	61	69	240
% of Station-SKUs	0.0%	1.3%	15.6%	6.7%	5.8%	9.4%	9.7%	9.1%	8.5%	6.4%	4.5%	4.3%	4.0%	3.8%	2.9%	1.9%	1.0%	1.1%	3.9%
Cumulative % of Station-SKUs	0.0%	1.3%	17.0%	23.7%	29.5%	38.9%	48.6%	57.7%	66.2%	72.6%	77.0%	81.3%	85.3%	89.0%	92.0%	93.9%	94.9%	96.0%	100.0%

Figure 7-5: Histogram of the Coefficient of Variation (CV) of inter-count times for 6,110 hospital Station-SKUs

7.3.2 The case of the Cardiac Catheterization lab

This particular ward is studied to illustrate how variability in inter-count times occurs. Summary statistics on inter-count times in the Cardiac Catherization Lab are presented in Table 7.3.

Row Labels	Number of SKUs	Average of (Mean Inter-Count Time)	Average of (Median Inter-Count Time)	Average of (CV of Inter-Count Time)	Standard Deviation of (CV of Inter-count times)
01.S_CL_A	513	7.57	5.82	0.78	0.34
01.S_CL_B	684	8.19	6.07	0.69	0.38
01.S_CL_C	448	10.93	6.01	0.90	0.62

Table 7.3: Summary statistics on the distribution of variability of inter-count times across different products in a station (one station-SKU = one observation)

Analysis of count transactions data for specific Station-SKUs showed groups of counts being performed together at each station for different products on Wednesday and Friday around 6:00 am, suggesting that a cycle-count program was implemented in this ward. This pattern is consistent for most products at the different stations of the Cardiac Catheterization Lab, and can be seen in Figure 7-6.

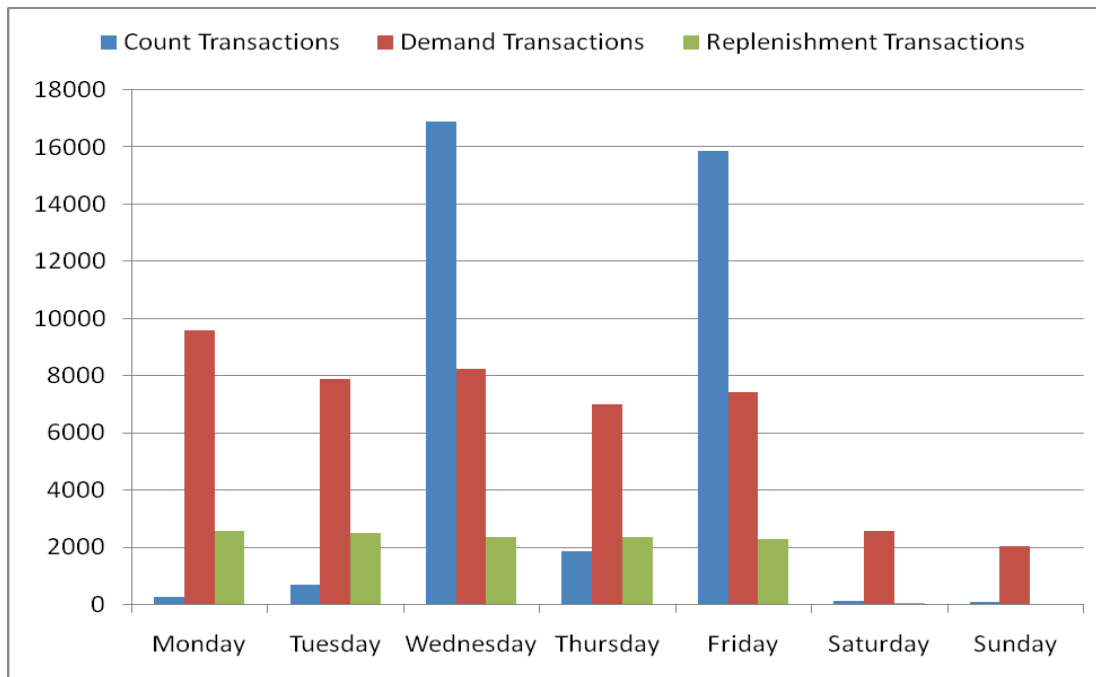


Figure 7-6: Total number of Count, Demand and Replenishment transactions by day of the week over the 15-month study period for Station A of the Cardiac Catheterization Lab⁵⁰

These results show a relatively lower level of variability in inter-count times in comparison with other wards at Lambda. Yet, the average of the coefficient of variation of inter-count times across SKUs stocked in Station A is 0.78, which is still high. If counts were always performed on schedule, the mean inter-count time would be 3.5 days and the coefficient of variation would be 0.61. The additional mean inter-count time and variation comes from products counted less fre-

⁵⁰ Highly similar patterns were found in Stations B and C, with counts predominantly taking place on Wednesday and Friday. The most notable difference was the absence of weekend activity in Stations B and C, as all weekend procedures were assigned to Room A.

quently, either because they are not part of the cycle-count program or because some of their counts were missed, which also explains the additional variability.

The Cardiac Catheterization Lab schedule does not ensure that counts are equally spaced in time, and the distribution of inter-count times has a mode at two days and another mode at five days. While the analytical model shows that this is suboptimal, this could potentially be an adjustment designed to reflect the nonstationarity of demand patterns throughout the week, and thus achieve counts performed after a set number of demand transactions (Iglehart & Morey, 1972). The demand and replenishment patterns throughout the week (Figure 7-6) show a reduced level of activity during the weekend. Considering that all counts happen at the beginning of the work day, the Wednesday morning count covers the activity from Friday to Tuesday, totaling 29,494 demand transactions annually, while the Friday morning count covers Wednesday and Thursday activity, which represents only 15,260 transactions annually. Thus, this example shows that the counts are unequally spaced even after adjusting for demand variations across the week. If the constraint of performing counts at the beginning of a weekday is kept, moving the Wednesday count to Tuesday morning would result in them being more equally spaced (see Table 7.4) both in terms of time and demand, and therefore would improve their effectiveness.

Count Day 1	Count Day 2	Mean Inter-Count Time	Theoretical Inter-Count Time C.V.	Mean Inter-Count Demand Transactions	Theoretical Inter-Count Demand C.V.*
Wednesday	Friday	3.5	0.61	339	0.45
Tuesday	Friday	3.5	0.20	339	0.05

Table 7.4: Illustrative Inter-count Time CV and Inter-count Demand CV for different Count Days, taking into account nonstationarity of demand during the week. This table assumes that counts are executed before the work day on those dates.

* Assuming deterministic demand following a weekly periodic pattern.

7.3.3 Testing for dynamic counting strategies

Interviews with materials managers revealed that large differences in cycle count frequency and variability existed among different products and wards, and that no formal cycle counting policy existed across the hospital. Materials managers attributed different levels of count activity to the particular technicians assigned to specific wards, and to the existence of ward-specific efforts (e.g., in the Cardiac Catheterization Lab). Moreover, the hospital had experimented with dynamic counting strategies. For instance, replenishment technicians were instructed to perform counts and correct inventory inaccuracies during replenishment transactions to the extent that their workload allowed it and they were not preventing nurses from accessing supplies. Because these activities were outside of the scope of a cycle counting program and the execution of a count during replenishment was not measured, such replenishment transactions were excluded from the analysis⁵¹.

In this subsection, whether the observed variability in inter-count times results from dynamic counting policies based on the book inventory state is investigated. It is possible that dynamic strategies could be based on other factors, in particular the actual inventory level⁵². However, physical inventory is only observed during inventory counts as opposed to during the entire period of the study, making it difficult to test for the presence of strategies based on physical inventory.

⁵¹ Because Counting On Replenishment (COR) is a dynamic counting policy, it violates the assumptions made in this model. However, sensitivity analyses of the Discrete-Event Simulation model showed that the detrimental effect of cycle count time variability on the service level still occurs when the COR policy is implemented.

⁵² An example of such a policy used by retailers in addition to cycle counting programs is the Zero Balance Walk. It consists of visually inspecting items with zero stock on hand and updating their records in the inventory system.

For each Station-SKU, the distribution of book inventory levels (a) as observed prior to each count was compared with (b) that as expected at a random point in time⁵³. Specifically, the null hypothesis is:

H₀: the relative proportions of the inventory levels are independent of whether they are measured during inventory counts or at a random point in time.

Some static cycle count programs (as opposed to dynamic counting strategies) could violate the null hypothesis, not because counts were triggered on the basis of the book inventory level, but rather as a result of a coincidence between the count scheduled time (e.g., Friday afternoons after the end of all operations) and particular levels of inventory given combinations of periodic demand patterns and replenishment schedules. Because of this broad definition of the null hypothesis, the procedure used is conservative in that correct rejection of the null does not necessarily indicate the presence of a dynamic inventory strategy. Since the expected values of the table cells were often below five, Fisher's exact test was used⁵⁴. Table 9.6 in Appendix H reports the results of hypothesis testing at the 0.01, 0.05 and 0.10 significance level in each station.

At the 0.05 significance level, the null hypothesis is rejected for 1,579 Station-SKUs (i.e., the null hypothesis fails to reject in 73.9% of the 6,110 Station-SKUs considered). This suggests that, at least for some products, hospital materials managers may be triggering counts on the basis of their book inventory level. A reasonable assumption is that such dynamic strategies may be

⁵³ For each level of book inventory, the number of days spent with book inventory equal to that level is rounded and constitutes the expected count, thus defining the empirical distribution of inventory levels across time. This distribution was obtained using transaction data for 5,855 station-SKUs among a total of 6,110.

⁵⁴ P-value estimates were obtained through Monte Carlo simulation (using 20,000 trials) by the `fisher.test()` command of the R statistical package.

followed in certain wards (and thus stations) and not in others. For each station with more than 50 station-SKUs, the proportion of SKUs for which the null is rejected at the 0.05 significance level is examined. In only eight out of 30 stations is this proportion greater than 50% (Table 7.5). In particular, in the Cardiac Catheterization ward (Stations S_CL_A, B, and C) described in the previous subsection, and where a cycle count program was described by materials managers during interviews, the null hypothesis is rejected for only 6%, 4% and 10% of Station-SKUs, respectively.

This analysis of dedicated count transactions suggests that a large fraction of products exists for which the variability in inter-count times is not attributable to dynamic counting strategies based on book inventory, but rather on suboptimal timing of cycle counts.

7.3.4 Estimating the count reduction potential at Lambda

The Count Inefficiency metric is defined in 7.2.2 as the percentage reduction in the number of counts that would keep the service level unchanged if the variability of inter-count times were eliminated for a single Station-SKU. This section extends this definition to groupings of Station-SKUs by applying the count inefficiency equation to calculate the absolute change in the number of counts⁵⁵ for each individual Station-SKU, and then aggregating the results.

Using this method on 6,110 Station-SKUs with more than ten counts, the total number of counts performed in the hospital during the study period could be reduced by 41.5% (i.e., from 261,934 to 153,165), if all variability were eliminated, or by 34.8% if each Station-SKU had an inter-

⁵⁵ i.e., a number of counts avoided, rounded to the lowest integer for conservativeness.

count time coefficient of variation of 0.25⁵⁶. These figures suggest a large improvement opportunity.

⁵⁶ These figures do not include the Station-SKUs for which the total number of counts performed was less than or equal to ten, which accounted for a total of 50,655 of the 312,589 counts performed (see Table 7.2). Conservatively assuming that the number of counts for these Station-SKUs is unchanged, the reduction would be 34.8% (target C.V. = 0.00) and 31.8% (target C.V. = 0.25).

Row Labels	Number of SKUs	Number of SKUs with p-Value < 0.05	% of SKUs for which null is rejected at 0.05 significance level
01.S_OR_SUT	129	114	88%
01.S_OR_LAP	66	57	86%
01.S_OR_GYN	93	79	85%
01.S_OR_LAP_2	104	73	70%
01.S_OR_GEN	92	55	60%
01.S_OR_URO	121	69	57%
01.S_OR_URO_2	52	29	56%
01.S_6C	197	106	54%
01.S_ENDO_5	51	24	47%
01.S_AMB_SURG	76	35	46%
01.S_OR_CAR_1	316	142	45%
01.S_ANES	122	51	42%
01.S_PACU	105	42	40%
01.S_5W	197	71	36%
01.S_ED_MINOR	73	20	27%
01.S_SICU_B_2	90	23	26%
01.S_OR_ORTD	51	13	25%
01.S_6W	170	37	22%
01.S_SICU_A_1	102	22	22%
01.S_SICU_B_1	99	19	19%
01.S_SICU_A_2	90	17	19%
01.S_CL_SR_1	412	73	18%
01.S_MICU_2	112	19	17%
01.S_ED_UTIL	114	19	17%
01.S_CCU_2	123	15	12%
01.S_7W	136	14	10%
01.S_CL_C	409	42	10%
01.S_7C	104	8	8%
01.S_CL_A	476	29	6%
01.S_CL_B	583	22	4%

Table 7.5: Proportion of products for which the null hypothesis is rejected at the 0.05 level (stations with less than 50 SKUs were excluded)

[Page intentionally left blank]

Chapter 8

Managerial Implications

The key issues for hospital materials managers faced with inventory inaccuracies are: (1) the identification of effective strategies to improve product availability and (2) the development of a holistic framework to implement and evolve these strategies over time.

This chapter builds on the simulation study to summarize the benefits of different strategies, and discusses their feasibility in a Lambda's hospital environment. An iterative framework for continuous improvement and the reduction of inventory inaccuracies is proposed and illustrated using a causal loop diagram.

8.1 Evaluation of Inventory Inaccuracy Strategies

This thesis considered four strategies for addressing the service level impact of inventory inaccuracy strategies, and their quantitative effects are reported in section 6.4: (1) Reducing Imperfect Demand Recording, (2) Counting More Frequently and Reducing Count Variability, (3) Counting on Replenishment, and (4) following a Dynamic Inventory Policy. The benefits and operational implications of each strategy are discussed below.

8.1.1 Reducing Imperfect Demand Recording

This strategy consists of conducting process improvement efforts to ensure that product usage is properly recorded by end-users. In essence, this is a prevention strategy in that it attempts to ad-

dress the root causes of inventory inaccuracy (DeHoratius, et al., 2008). This strategy was the one that Lambda pursued when this study was initiated. This strategy was being followed by the hospital for multiple reasons.

First, achieving high Recording Accuracy guarantees that the product usage data can be accurately tied to a particular patient and ward, and therefore enables billing integration. While beyond the direct scope of this thesis, the use of these data for billing integration is desirable because it enables the hospital to understand its cost structure with a high level of granularity (i.e., down to the patient, procedure and physician level). When the patient is insured, the hospital is paid a fixed fee per procedure. Therefore, detailed supply cost information is necessary to measure the profitability of different procedures. When the patient is uninsured, products above a critical threshold (e.g., \$20) are billed to the patient's account, and therefore accurate information is desirable. When this research was initiated with Lambda, one of the goals of implementing ADMs was to reduce inaccuracies to an acceptable level so that billing integration could be implemented. At the time of this writing, three and a half years after the ADM deployment began at Lambda, billing integration had not been activated because of Imperfect Demand Recording.

Second, vendors of Automated Dispensing Machines recommend that hospitals focus on prevention strategies and staff accountability to achieve high Recording Accuracy. However, process improvement and accountability are difficult to implement without detailed knowledge of the sources of inaccuracies, including the stations and the individuals responsible for them. Because the value proposition of ADMs rests heavily on the cost-saving potential of automation (Owens & Minor, 2005), and frequent inventory counting is typical when ADMs are not deployed, correction strategies were often perceived by hospital staff as a failure of the ADM system, and were not advocated by the vendor. This is because inventory counting is labor-intensive and erodes the labor savings that the vendor claims are achievable with ADM systems. Furthermore, introducing counting strategies constitutes a return to the prior process, as was suggested by a hospital manager during a visit: *"If we still have to perform inventory counts, why do we need Automated Dispensing Machines?"*

In summary, achieving and sustaining a high Recording Accuracy should be the medium- to long-term goal of the hospital managers to get the full benefits of Automated Dispensing Machines.

8.1.2 Counting more frequently and reducing Count Variability

Correcting the inventory record more frequently through physical audits improves service levels significantly⁵⁷ and provides improved estimates of the recording accuracy⁵⁸. This strategy is therefore important in that it offers immediate and large service level improvements when the recording accuracy is low, and provides data to assist process improvement efforts. However, conducting more inventory counts requires additional labor and therefore adds costs for the hospital.

On the other hand, reducing cycle count variability improves the service level and has no impact on costs because the count frequency is unchanged. This finding offers an opportunity to maintain the current service level while reducing the number of counts by at least 31.8% across the entire hospital, and/or to achieve higher service levels without increasing costs⁵⁹.

At the scale of the hospital, this leads to an opportunity to reallocate counts freed by variability reduction to products that are currently not counted frequently enough.

⁵⁷ See Table 6.5: Physical and Book View In-stock Probability for different values of the Recording Accuracy and Count Periodicity.

⁵⁸ See section 5.4.1: Challenges to estimating Recording Accuracy from ADM data.

⁵⁹ See section 7.2.2 for the quantification of the possible reduction in counts and section 7.3.4 for its application to empirical hospital data.

8.1.3 Counting on Replenishment

The simulation model demonstrated that Counting on Replenishment increases product availability by increasing the total number of counts performed and removes the overly optimistic bias in the Book Stock-Out Frequency metric⁶⁰, which is reported by Automated Dispensing Machines and was tracked by hospital managers throughout the study. This strategy was implemented by the hospital halfway through the study at the suggestion of the ADM vendor's consulting staff, and prior to receiving knowledge of the results of this study.

However, Counting on Replenishment suffers from severe operational drawbacks. First, it concentrates the workload of materials management technicians by adding a process step to the replenishment, thus slowing down delivery of product to the stations in the morning. At Lambda, this resulted in increased complaints by end-users of not being able to access the machines as they were being replenished, as well as low labor utilization in the early afternoon. Second, contrary to a dedicated count transaction, it is impossible to monitor whether a count on replenishment truly has been performed and to know with certainty the time of the last count. Finally, the simulation shows that counting on replenishment is less efficient than periodic counts; for a given total count budget, periodic counts alone achieve a higher service level. Intuitively, this can be understood easily, as counts performed during replenishment are less likely to prevent a stock out because they occur *after* the ordering decision has been made.

For these reasons, alternative counting policies should be pursued in lieu of Counting on Replenishment.

⁶⁰ The bias is removed if counts are performed at every replenishment, and is only removed in steady state, i.e. the stock outs are usually detected with a variable delay after which they took place.

8.1.4 Dynamic Inventory Policy

The dynamic inventory policy presented in this thesis is an example of an integration strategy. It takes imperfect demand recording into account during the replenishment decision. While other integration strategies are more sophisticated in their theoretical grounding, the Dynamic Inventory Policy heuristic developed in this thesis has the advantage of being computationally efficient and self-tuning. It updates itself as new information about inventory inaccuracy is obtained during audits, thus avoiding the need for one-time parameter estimation. For instance, if Recording Accuracy drops, the heuristic detects this after the subsequent count and attempts to compensate for the drop without managerial intervention.

Strategies that integrate inventory inaccuracy into inventory policy are very recent (DeHoratius, et al., 2008; Kok & Shang, 2007). That inventory managers may not yet readily understand them, as well as the need for sophisticated software, has so far limited their real-world applications. Furthermore, hospital materials managers are not able to experiment with inventory policy beyond adjusting its parameters because the ADM vendor designs and maintains the software. This presents both a challenge and a rare strategic opportunity, as it makes piloting new initiatives more difficult because of the need for the involvement of the ADM vendor. However, it is a significant leverage point for facilitating the adoption of improved inventory policies. While the hospital industry is particularly fragmented, and disseminating innovations in healthcare is difficult (D. Berwick, 2003), only three main ADM vendors⁶¹ deploy and maintain ADM software in thousands of hospitals, which means that software innovations could potentially spread rapidly.

⁶¹ See Supply Chain Automation Vendors, section 2.2.2: Stakeholders, page 35.

8.2 A Holistic and Iterative Framework for Improving Product Availability in the Hospital

In the previous section, the benefits and drawbacks of different improvement strategies were discussed. This section proposes an iterative framework that takes into account the interdependency among these strategies. First, counting more often and more consistently provides more precise estimates of the Recording Accuracy, which are instrumental in identifying corrective actions and providing feedback to end-users to reduce Imperfect Demand Recording. Second, counting policies affect the bias and variance of service level estimates, which can become severely inaccurate and give managers a false sense of security.

For this reason, this thesis recommends that hospital managers avoid using current service level metrics to track the evolution of process improvement efforts, and rely instead on metrics that reflect the root cause of the problem, such as Recording Accuracy.

8.2.1 Framework description

This thesis proposes to initially focus on correction strategies prior to implementing prevention or integration strategies. This is consistent with the Statistical Process Control (SPC) literature, which recommends standardizing existing processes, defining precisely characteristics to be improved, and collecting data and tracking these characteristics prior to conducting improvement activities (Carey & Lloyd, 1995). The proposed framework applies this principle to inventory inaccuracies and combines Correction, Prevention, and Integration strategies in a sequential and iterative manner, as follows.

1. Correction

- a. Setting a target cycle count frequency. For each station, define a cycle counting schedule for the different SKUs based on the current Recording Accuracy levels. Several steps can facilitate a practical schedule:

- i. Dedicating equally spaced time slots to cycle counts chosen at periods of low machine and labor utilization;
 - ii. Choosing a limited number of cycle count frequencies, e.g., bi-weekly, weekly, bi-monthly, and monthly frequencies;
 - iii. Adopting a computerized system to automatically generate lists of items that need to be counted each day.
- b. Counting consistently
- i. Using control charts to monitor the mean and standard deviation of inter-count times;
 - ii. Actively investigating missing or ill-timed counts.

2. Prevention

- a. Set a target Recording Accuracy for each station.
- b. Reporting estimates of Recording Accuracy to end-users:
 - i. At frequent intervals, but only when the estimates are statistically valid⁶²;
 - ii. Separate different groups of products, e.g., Closed Door vs. Open Shelf;
 - iii. Consider reporting Recording Accuracy on a modified scale consistent with the Defect Per Million convention to make the differences in Recording Accuracy meaningful to end-users (see section 5.3, The Role of Framing, page 76 for a discussion of this issue).
- c. Conduct training sessions focused on problem areas to resolve technical and workflow problems (Loop B1 on Figure 8-1)

⁶² More frequent estimates are based on fewer counts and therefore have higher variance. More frequent and consistent counts reduce the variance of recording accuracy estimates.

- d. **Iterate** by adjusting the target count frequency as Recording Accuracy evolves (Loop B2 on Figure 8-1)

3. Integration

Once robust counting processes and higher Recording Accuracy levels are achieved, dynamic ordering policies should be considered to improve the performance of the supply chain.

The Correction and Integration steps can be summarized in the following Causal Loop Diagram:

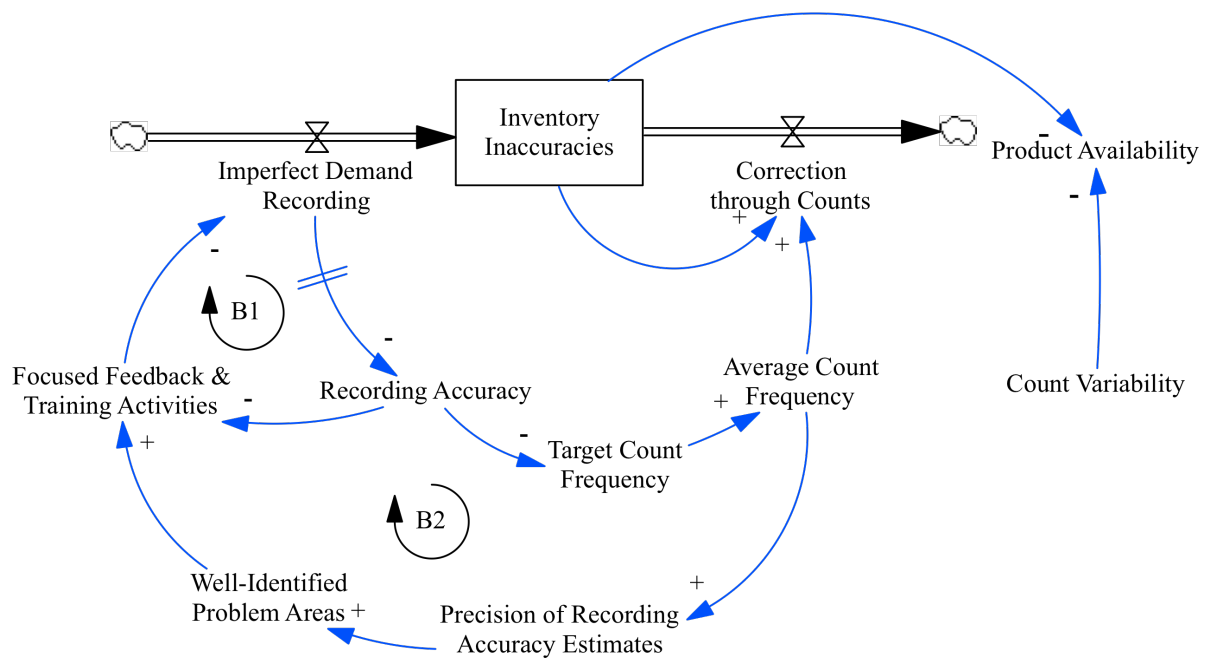


Figure 8-1: Causal Loop Diagram of the Iterative Improvement Framework

The double slash indicates the delay between Imperfect Demand Recording and the Recording Accuracy estimate.

8.2.2 Discussion

Contrary to intuition of eliminating directly the root causes, this framework starts with Correction strategies. In addition to the statistical process control analogy, specific behavioral and organizational reasons motivate this order.

To the casual observer, Imperfect Demand Recording is a problem that can be solved simply through accountability, by defining it as an inappropriate behavior for which repeat offenders suffer harsh consequences. The literature on human factors has found this approach ineffective in many industries because a culture of blame prevents error reporting and improvement activities. This insight is being increasingly recognized in health care, including in the area of medication errors (D. M. Berwick, 1998; Fortescue, et al., 2003; Kaushal, et al., 2001; Leape, 2009). Furthermore, compliance is practically difficult to enforce when Imperfect Demand Recording is widespread in a ward.

On the other hand, achieving process control in correction activities poses fewer difficulties than for Imperfect Demand Recording, for the following reasons: (1) materials management technicians are under the authority of a single manager for whom supply chain performance is a key objective, (2) the number of materials management technicians is small relative to the number of end-users (25 FTEs relative to 2,300 at Lambda), and (3) Automated Dispensing Machines track counting behavior at the individual level.

Finally, it is important to note that counting consistently delivers a potential Pareto improvement in the form of service level increases with little to no incremental costs⁶³. This creates an oppor-

⁶³ When count frequency, which is the main driver of counting labor costs, is kept constant. Scheduling cycle counts may require additional resources but off-the-shelf software solutions should make this cost negligible.

tunity for hospital managers to optimize the allocation of their count budget across stations and products, as counts are no longer necessary because their variability has been reduced and can be allocated to products needing a higher count frequency, or can be eliminated altogether. Avoiding increases in the overall counting workload of materials management technicians is important for two reasons: (1) it represents a cost to the hospital, and (2) increasing count frequency in a warehousing context resulted in *decreased* inventory accuracy because it is a tedious task and technicians did not truly execute the recommended counts, using the book inventory as a “maximum likelihood estimate” for the physical inventory (Millet, 1994).

Finally, interviews with materials management technicians and end-users revealed that dynamic decisions could be a source of confusion. This observation fits within the literature on the value of consistency in decision-making (Bowman, 1963), which notes that consistent managerial decisions over time can be more beneficial than explicit solutions to cost-models when intangibles such as stock out costs have to be estimated or assumed. When their variability is reduced, periodic cycle counts satisfy this consistency property.

Chapter 9

Conclusions & Contributions

This chapter summarizes the conclusions reached in this research, starting with the characterization of inventory inaccuracies in the hospital. The insights derived from the discrete-event simulation model on the effects of Imperfect Demand Recording are presented, followed by their implications for hospital inventory management and a discussion of the socio-technical aspects of the inventory inaccuracy problem.

The contributions of this research to the hospital system domain, inventory management, and Engineering Systems are then presented, as well as suggestions for further research.

9.1 Conclusions

9.1.1 Characterization of Inventory Inaccuracies in the Hospital

Analysis of Automated Dispensing Machine transaction data from a 300-bed New England hospital (Lambda) showed that inventory inaccuracy is a significant issue in the hospital context. A typology of the sources of inventory inaccuracies was constructed and Imperfect Demand Re-

ording, i.e., the act of not properly recording product usage, was identified as a key source of inventory inaccuracies.

Existing metrics for inventory inaccuracy reported by Automated Dispensing Machines suffer from several problems that reduce their usefulness to managers: (1) they are affected by the number of counts performed in a given period, and therefore do not reflect the underlying behavior of end-users, (2) they are skewed toward fast-moving products at the expense of slow-moving products, and (3) they are not robust to outlier observations.

For these reasons, this thesis proposes Recording Accuracy as a new metric of Imperfect Demand Recording and provides methodology to estimate it from Automated Dispensing Machines count data. Its application to Lambda's records shows that depending on the ward, between five and thirty percent of product usage is not recorded, i.e., Recording Accuracy ranges between 70% and 95%.

9.1.2 Drivers of Product Availability

A discrete-event simulation model showed a large and statistically significant detrimental effect of Imperfect Demand Recording on product availability, which increased with increasing average time between inventory counts. For instance, assuming counts occur every three weeks, as Recording Accuracy drops from 100% to 85%, the fill rate declines from 98.2% to 89.4%. This same simulation scenario found that variability in the time between counts further reduces the fill rate by 5% to 15%. An analytical model confirmed that variability in cycle counting undermines the effectiveness of inventory counts under more general assumptions about inventory inaccuracies, and this effect was further quantified.

9.1.3 Bias and Inadequacy of Existing Service Level Metrics

Existing metrics of product availability available from Automated Dispensing Machine data were found to suffer from serious limitations.

First, the stock-out frequency metric, which is reported by ADMs and used by hospital managers, does not account for stock-out duration. Therefore, it does not reflect the experience of end-users. The stock out frequency is not monotonic in the fill rate, as when the fill rate decreases to very low levels, the stock out frequency reaches a plateau or even decreases because items that are out of stock cannot go out of stock again. This first observation is general regardless of whether low service levels are caused by inventory inaccuracy or insufficient safety stock.

Second, unless a count is performed immediately prior to the replenishment delivery, the book stock out frequency metric (calculated from ADM data) is biased and overly optimistic. For instance, using the previous simulation scenario, the book stock out frequency is equal to 0.32 stock out events per week when the physical (i.e., correct) stock out frequency is equal to 0.11 stock out events per week. Moreover, the book view stock out frequency metric is not monotonic in the physical (i.e., true) stock out frequency. Finally, the book in-stock probability suffers from the same bias phenomenon.

9.1.4 Implications for Hospital Inventory Management

The previous findings suggest that managers should avoid using current service level metrics reported by Automatic Dispensing Machines, since those metrics: (1) do not reflect the true service level experienced by caregivers; (2) are not a reliable proxy or performance indicator during improvement activities, as the metric can improve while the true situation is deteriorating.

Instead, hospital managers should focus on measuring the root causes of inventory inaccuracies to drive their process improvement efforts. A sequential and iterative framework for improving product availability is proposed, as follows: (1) start by setting targets for the frequency of inventory counts and increasing their effectiveness by reducing their variability, (2) obtain improved estimates of Recording Accuracy and use them to focus training and process improvement efforts on problematic areas, and (3) iterate by adjusting the target count frequency as Recording

Accuracy evolves, creating a balancing loop that gradually reduces the resources devoted to inventory counts.

9.2 Contributions

This study considered the hospital supply chain as an engineering system and investigated the role of inventory inaccuracies in its performance. Using a multimethod (i.e., combining interviews, econometric analysis of real-world data, discrete-event simulation, and analytical modeling) and multilevel (i.e., hospital, ward, and item levels) approach proved beneficial throughout this research. The results most relevant to the literature on inventory inaccuracies and variability are discussed first, followed by remarks pertinent to engineering systems.

First, variability in the time between inventory cycle counts is empirically established in hospital Automated Dispensing Machine transaction data. Under the assumption of independence of count timing with the state of the inventory (which was tested using empirical data), this variability property of the system is shown to reduce service levels. From a theoretical perspective, a distribution-free approximation of the cost of variability was derived and was found to be proportional to a nonlinear S-curve. Practically, this implies that managers can improve service levels at no cost and/or reduce counting costs by ensuring that cycle counts are executed according to an equally spaced schedule.

Second, hospital managers were found to rely on metrics from Automated Dispensing Machines that are misaligned with the experience of end-users and are biased by Imperfect Demand Recording. This problematic situation is not accidental. It results from metrics and technical systems designed with incorrect assumptions about the compliance of individuals with prescribed processes. A sequential and iterative framework is proposed, which seeks to improve performance by taking a holistic approach using technology to detect changes in process compliance. These findings were presented at Lambda hospital at the end of this study and led to a shift in focus from biased service level metrics toward counting policies.

9.3 Suggestions for Future Research

Three areas for further research stem from this thesis:

- Field experiments evaluating interventions such as the ones proposed in this thesis, including Prevention, Correction, and Integration strategies. However, gaining access to management and data, enforcing appropriate experimental controls, and ensuring external validity through the use of multiple sites make this a challenging task.
- Theoretical and applied research on unbiased estimation methods for service level metrics, such as the in-stock probability or the fill rate in the presence of inventory inaccuracies.
- More generally, empirical and theoretical research characterizing how human behavior deviates from assumptions embedded in socio-technical systems and evaluating the effects of such deviations on system performance to identify improvement opportunities.

Bibliography

ACHE (2009). American College of Healthcare Executives Announces "Top Issues Confronting Hospitals: 2008". Retrieved from http://www.ache.org/Pubs/Releases/2009/CEOTopIssues_2008.pdf

Agnew, R. (1972). Inequalities with application in economic risk analysis. *Journal of Applied Probability*, 441-444.

Agrawal, N., & Smith, S. A. (1996). Estimating negative binomial demand for retail inventory management with unobservable lost sales. *Naval research logistics*, 43(6), 839--861.

AHA, & Avalere (2008). *Trendwatch Charbook 2008*.

AHA, & Levin Group (2006). *Trendwatch Charbook 2006*.

American College of Healthcare Executives (ACHE) (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century; Advance Copy, Uncorrected Proofs*: National Acad. Press.

Amos, D. (1974). Computation of modified Bessel functions and their ratios. *Mathematics of Computation*, 239-251.

Atali, A., Lee, H., & Ozer, O. (2006). If the inventory manager knew: Value of visibility and RFID under imperfect inventory information. [Working Paper]. *Working Paper*.

Barrett, G. (2009, March 11, 2009). "New" Cardinal Health (Post-Spin). Paper presented at the Barclays Capital Healthcare Conference.

- Bartolomei, J., Hastings, D., de Neufville, R., & Rhodes, D. (2006). *Screening for Real Options" In" an Engineering System: A Step Towards Flexible System Development--PART I: The Use of Design Matrices to Create an End-to-End Representation of a Complex Socio-Technical System.*
- Beier, F. J. (1995). THE MANAGEMENT OF THE SUPPLY CHAIN FOR HOSPITAL PHARMACIES: A FOCUS ON INVENTORY MANAGEMENT PRACTICES. *Journal of Business Logistics*, 16(2), 153-173.
- Berndt, E. R. (2007). Anatomy of the U.S. Health Care System. In M. S. C. J. Notes (Ed.).
- Berwick, D. (2003). Disseminating innovations in health care. *Jama*, 289(15), 1969.
- Berwick, D. M. (1998). Crossing the boundary: changing mental models in the service of improvement. *International Journal for Quality in Health Care*, 10(5), 435-441.
- Biviano, M., Tise, S., & Dall, T. (2007). What is Behind HRSA's Projected US Supply, Demand, and Shortages of Registered Nurses. *Modelling Our Future: Population Ageing, Health and Aged Care*, 343.
- BLS (2010). Occupational Outlook Handbook. from US Department of Labor, Bureau of Labor Statistics: <http://www.bls.gov/oco/ocos083.htm>
- Bowman, E. H. (1963). CONSISTENCY AND OPTIMALITY IN MANAGERIAL DECISION-MAKING. [Article]. *Management Science*, 9(2), 310-321.
- Britt, R. (2007). Growing share of 'Big Three' gets federal attention; Giant wholesalers dominate market, to hit quarter-trillion mark in sales. *MarketWatch*, (May 30, 2007). Retrieved from "<http://www.marketwatch.com/news/story/growing-share-big-three-drug/story.aspx?guid={9F3862C1-7E3A-4D82-94B7-A8E4AE5DAE6A}>"
- Brooks, R., & Wilson, L. (2007). *Inventory record accuracy: unleashing the power of cycle counting*: Wiley.

- Burns, L. R. (2002). *The health care value chain: producers, purchasers, and providers*: Jossey-Bass.
- Carey, R., & Lloyd, R. (1995). *Measuring quality improvement in healthcare: A guide to statistical process control applications*: American Society for Quality.
- CMS (2009). National Health Expenditures by type of service and source of funds, CY 1960-2007.
- Cox, D. R. (1962). *Renewal theory*. London; New York: Methuen; Wiley.
- CSA (2001). INVENTORY MANAGEMENT. *Chain Store Age*, 77(12), 3A.
- CSA (2003). INVENTORY MANAGEMENT 2003: COMMUNICATION AND TECHNOLOGY BEAR RENEWED PROMISE. *Chain Store Age*, 79, 3A-11A.
- Davies, R. (1973). Numerical inversion of a characteristic function. *Biometrika*, 60(2), 415-417.
- de Kok, A. G., van Donselaar, K. H., & van Woensel, T. (2008). A break-even analysis of RFID technology for inventory sensitive to shrinkage. *International Journal of Production Economics*, 112(2), 521-531.
- DeHoratius, N., Mersereau, A. J., & Schrage, L. (2008). Retail inventory management when records are inaccurate. *Manufacturing & Service Operations Management*, 10(2), 257.
- DeHoratius, N., & Raman, A. (2008). Inventory record inaccuracy: An empirical analysis. *Management Science*, 54(4), 627-641.
- DeScioli, D. T. (2005). *Differentiating the Hospital Supply Chain For Enhanced Performance*. Unpublished M.Eng. Thesis, Massachusetts Institute of Technology.

- Ennis, D. M., & Bi, J. (1998). THE BETA-BINOMIAL MODEL: ACCOUNTING FOR INTER-TRIAL VARIATION IN REPLICATED DIFFERENCE AND PREFERENCE TESTS. *Journal of Sensory Studies*, 13(4), 389-412.
- Everson, P. J., & Bradlow, E. T. (2002). Bayesian Inference for the Beta-Binomial Distribution via Polynomial Expansions. *Journal of Computational & Graphical Statistics*, 11(1), 202--207.
- FDA (2009). U.S. Food and Drug Administration, About FDA, What We Do, 2009, from <http://www.fda.gov/AboutFDA/WhatWeDo/default.htm>
- Federal Food, Drug, and Cosmetic Act, as amended § 201 (1938).
- Finkelstein, S. (2003). *Cardinal Health and Allegiance*: Tuck School of Business at Dartmouth.
- Fleisch, E., & Tellkamp, C. (2005). Inventory inaccuracy and supply chain performance: a simulation study of a retail supply chain. *International Journal of Production Economics*, 95(3), 373-385.
- Fortescue, E., Kaushal, R., Landrigan, C., McKenna, K., Clapp, M., Federico, F., et al. (2003). Prioritizing strategies for preventing medication errors and adverse drug events in pediatric inpatients. *Pediatrics*, 111(4), 722.
- Foundation, N. R. F., & BearingPoint (2005). *Retail horizons : benchmarks for 2005, forecasts for 2006*. [McClean, VA]; [Washington, D.C.]: BearingPoint ; NRF Foundation.
- Gentry, C. R. (2005). Theft and Terrorism. *Chain Store Age*, 81(7), 66-68.
- Gold, M. (1996). *Cost-effectiveness in health and medicine*: Oxford University Press, USA.
- Griffiths, D. A. (1973). Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics*, 29(4), 637--648.

- Hadley, G., & Whitin, T. (1961). A family of inventory models. *Management Science*, 351-371.
- Hagland, M. (2005). Supply traffic control. A consolidated materials management system enhances savings, processes and time for patient care. *Healthcare Informatics*, 22(9), 43-44.
- Handbook, O. (2007). US Department of Labor, Bureau of Labor Statistics. *Bulletin*.
- Hollinger, R. C. (2003). National Retail Security Survey. Retrieved November 20, 2009: <http://www.jrrobertssecurity.com/security-news/security-crime-news0024.htm>
- Iglehart, D. L., & Morey, R. C. (1972). Inventory systems with imperfect asset information. *Management Science*, 18(8), 388-394.
- Institute of Medicine, C. o. t. C. o. U. (2004). *Insuring America's health: principles and recommendations*: National Academies Press.
- Jayaraman, V., Burnett, C., & Frank, D. (2000). Separating inventory flows in the materials management department of Hancock Medical Center. *Interfaces*, 56-64.
- Kang, Y., & Gershwin, S. B. (2005). Information inaccuracy in inventory systems: stock loss and stockout. *IIE Transactions*, 37(9), 843-859.
- Karlis, D., & Ntzoufras, I. (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine*, 25(11).
- Kaufman, L., & Rousseeuw, P. (1990). Finding groups in data. An introduction to cluster analysis.
- Kaushal, R., Bates, D., Landrigan, C., McKenna, K., Clapp, M., Federico, F., et al. (2001). Medication errors and adverse drug events in pediatric inpatients. *JAMA: The Journal of the American Medical Association*, 285(16), 2114.

- Keeney, R., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*: Cambridge University Press.
- Klibanov, O. M., & Eckel, S. F. (2003). Effects of automated dispensing on inventory control, billing, workload, and potential for medication errors. *American Journal of Health-System Pharmacy*, 60(6), 569.
- Kok, A., & Shang, K. (2007). Inspection and Replenishment Policies for Systems with Inventory Record Inaccuracy. *Manufacturing & Service Operations Management*, 9(2), 185.
- Kok, A., & Shang, K. (2009). Evaluation of Supply Chains with Inventory Record Inaccuracy and Implications on IT Investments. Duke Fuqua School of Business.
- Kumar, S., & Arora, S. (1992). Effects of inventory miscount and non-inclusion of lead time variability on inventory system performance. *IIE Transactions*, 24(2), 96-103.
- Lapierre, S. D., & Ruiz, A. B. (2007). Scheduling logistic activities to improve hospital supply systems. *Computers & Operations Research*, 34(3), 624-641.
- Law, A. M. (2007). *Simulation modeling and analysis* (4th ed.). Boston: McGraw-Hill Higher Education.
- Leape, L. (2009). Errors in medicine. *Clinica Chimica Acta*, 404(1), 2-5.
- Lee, H., & Özer, Ö. (2007). Unlocking the Value of RFID. *Production and Operations Management*, 16(1), 40-64.
- Lee, J. C., & Sabavala, D. J. (1987). Bayesian Estimation and Prediction for the Beta-Binomial Model. *Journal of Business & Economic Statistics*, 5(3), 357--367.
- MacCrimmon, K. (1969). Improving the System Design and Evaluation Process by the Use of Trade-Off Information: An Application to Northeast Corridor Transportation Planning.

- Martin, Z. (2006). Supply chain mantra:'never run out'. *Health data management*, 14(10), 70.
- Medline (2009). Medline Signs Prime Vendor Agreement With Prime Healthcare Services. Retrieved from <http://www.marketwire.com/press-release/Medline-Industries-Inc-962578.html>
- Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., et al. (2001). Defining supply chain management. *Journal of Business logistics*, 22(2), 1-26.
- Millet, I. (1994). A Novena to Saint Anthony, or How to Find Inventory by Not Looking. *Interfaces*, 24(2), 69.
- Morey, R. C. (1985). Estimating service level impacts from changes in cycle count, buffer stock, or corrective action. *Journal of Operations Management*, 5(4), 411-418.
- Morey, R. C., & Dittman, D. A. (1986). Optimal timing of account audits in internal control. *Management Science*, 32(3), 272-282.
- Muller, K. (2004). Advanced systems simulation capabilities in SimPy. *Europython 2004*.
- Neil, R. (2003). Automatically easier. Supply chain benefits from point-of-pull. *Materials Management in Health Care*, 12(9), 18.
- Nicholson, L., Vakharia, A. J., & Erenguc, S. S. (2004). Outsourcing inventory management decisions in healthcare: Models and application. *European Journal of Operational Research*, 154(1), 271-290.
- OECD (2009). Health at a Glance 2009. Retrieved December 18, 2009: <http://www.oecd.org/health/healthataglance>
- Oliveira, J. (2005). *Enabling long term value added partnership in the healthcare industry*. Massachusetts Institute of Technology.

- Opolon, D., Schumacher, K., & Sheffi, Y. (2009). *Impact of Imperfect Demand Recording on Service Levels: the Case of Hospital Supplies*. Paper presented at the POMS 20th Annual Conference.
- Owens, & Minor (2005). QSight Clinical Inventory Control. Retrieved from <http://www.owens-minor.com/omsolutions/QSightBrochureSept2005.pdf>
- Pedersen, C. A., Schneider, P. J., & Scheckelhoff, D. J. (2006). ASHP national survey of pharmacy practice in hospital settings: Dispensing and administration-2005. *American Journal of Health-System Pharmacy*, 63(4), 327.
- Piasecki, D. (2003). *Inventory accuracy: people, processes, & technology*: Ops Publishing.
- Prentice, R. L. (1986). Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *Journal of the American Statistical Association*, 81(394), 321--327.
- Raman, A. (2000). Retail-data quality: evidence, causes, costs, and fixes. *Technology in Society*, 22(1), 97-109.
- Redman, T. C. (1995). Improve Data Quality for Competitive Advantage. *Sloan management review*, 36(2), 99.
- Reid, P. (2005). *Building a better delivery system: a new engineering/health care partnership*: National Academy Press.
- Riddalls, C. E., Bennett, S., & Tipi, N. S. (2000). Modelling the dynamics of supply chains. *International Journal of Systems Science*, 31(8), 969-976.
- Rinehart, R. F. (1960). Effects and Causes of Discrepancies in Supply Operations. *Operations Research*, 8(4), 543-564.

- Rivard-Royer, H., Landry, S., & Beaulieu, M. (2002). Hybrid stockless: a case study - Lessons for health-care supply chain integration. *International Journal of Operations & Production Management*, 22(4), 412-424.
- Rossetti, M., Collins, T., & Kurgund, R. (2001). *Inventory Cycle Counting—A Review*.
- Schneller, E. S. (2009). *Meeting the Need for Strategic Savings: Health Care Sector Advances*, Inc.
- Schneller, E. S., & Smeltzer, L. R. (2006). *Strategic Management of the Health Care Supply Chain*: Jossey-Bass.
- Schneller, E. S., & Smeltzer, L. R. (2006). *Strategic Management of the Health Care Supply Chain*: Jossey-Bass.
- Schrady, D. A. (1970). Operational definitions of inventory record accuracy. *Naval research logistics quarterly*, 17(1).
- Sheppard, G. M., & Brown, K. A. (1993). Predicting inventory record-keeping errors with discriminant analysis: a field experiment. *International journal of production economics*, 32(1), 39-51.
- Singh, M., & Rice, J. B. (2006). Transforming the Healthcare Supply Chain: Roundtable Proceedings. Massachusetts Institute of Technology, MIT Center for Transportation and Logistics.
- Skellam, J. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society*, 296-296.
- Sussman, L., & Gupta, Y. (1992). Different perspectives of inventory management among physicians and hospital administrators. *American journal of surgery*, 164(1), 1.
- Team, R. D. C. (2004). R: A language and environment for statistical computing.

- Tenet (2009). Tenet, Company Fact Sheet Retrieved March 25, 2009, from <http://www.tenethealth.com/TenetHealth/OurCompany/CompanyFacts>
- Tucker, A. (2004). The impact of operational failures on hospital nurses and their patients. *Journal of Operations Management*, 22(2), 151-169.
- Tucker, A., & Spear, S. (2006). Operational failures and interruptions in hospital nursing. *Health Services Research*, 41(3 Pt 1), 643.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453.
- VA (2009). Fact Sheet: Facts About the Department of Veterans Affairs Retrieved March 25, 2009, from <http://www1.va.gov/opa/fact/vafacts.asp>
- Wikipedia (2010). Medoid Article. Retrieved from <http://en.wikipedia.org/wiki/Medoid>
- Williams, D. A. (1975). The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics*, 31(4), 949--952.
- Yee, T. (2008). The VGAM Package. *The Newsletter of the R Project*, 8(2), 28.
- Zipkin, P. (1986). Inventory service-level measures: convexity and approximation. *Management Science*, 975-981.

Appendix A. Interview list

<i>Title</i>	<i>Organization</i>
Director of Purchasing	Lambda
Director of Materials Management	Lambda
VP of Supply Chain	Lambda
Nursing Manager	Lambda
Replenishment Technician 1	Lambda
Replenishment Technician 2	Lambda
OR Materials Support	Lambda
Program Manager*	Distributor and Lambda
Technology Staff Specialist	Boston-area hospital 1
Director of Perioperative Services	Boston-area hospital 2
Director of Pharmacy	Boston-area hospital 3
Capital Equipment Purchaser	Hopital Pitié-Salpêtrière, Paris, France
Director of Pharmacy (includes sterile medical devices)	Institut Gustave Roussy, Villejuif, France
Director, Global Supply Chain Processes Medical Devices & Diagnostics Group	Medical Device Manufacturer
Regional Sales Manager	Medical Device Manufacturer

* Title altered to preserve confidentiality.

Appendix B. Investigation of Demand Autocorrelation

The question of the presence of auto correlation in demand data is a recurring one in the inventory control literature. While most authors generally accept this assumption in theoretical papers, it is still important to examine its validity in the context of the hospital. Potential sources of auto correlation of demand include:

- Periodic patterns: there are likely to be periodic (daily, weekly and possibly annual) patterns in the demand data. For instance, the activity of many wards of an hospital is typically lower during weekends. These patterns are a common occurrence and typically manifest themselves by an Auto Correlation Function (ACF) plot oscillating around zero. The traditional procedure is to deseasonalize the data, but for slow-moving items this is a difficult task due to a large fraction of zero observations.
- Grouping of similar procedures through scheduling: many hospitals typically schedule similar procedures (for instance in an operating room) together in order to increase operational efficiency and safety. This may make the demand for certain operating-room related products to have positive serial correlation within the day.
- Recurring use: some products (for instance pharmaceuticals) may be used as part of the treatment regimen for a particular patient. This can create some auto correlation, particularly for slow-moving items where the impact of a single patient may be more important.

Auto correlation testing procedure

To test for auto correlation in the data, we group the historical demand data in periods of length T , where T is expressed in days. In order to avoid taking into account products taken by mistake, the auto correlation test is conducted on the net sales (i.e. sales - returns during the period). Our procedure is then the following:

- a) For each product in a particular location, plot the Auto Correlation Function with lags of up to 14 days (two weeks).
- b) Inspect the graph for patterns of positive serial correlation.
- c) If the data has a periodic pattern, we can fit a seasonal model and conduct the test on the residuals of the model, making seasonal effects manageable by the model⁶⁴.
- d) Conduct a Ljung-Box test with lags of up to 14 days. We choose the Ljung-Box test over the traditional Box-Pierce test because it has better performance for all sample sizes, particularly in small samples where the Box-Pierce test has poor performance.

Adapted from Wikipedia, “Ljung-Box test”:

The Ljung-Box test tests whether any group of autocorrelations of a time series are different from zero. Because it tests the randomness of the time series based on a number of lags, rather than at a specific lag, it is a portmanteau test. The null hypothesis is that the data are random, and the test statistic is:

⁶⁴ The Demand model with probabilistic recording is extended to the case when demand rates vary periodically in Appendix C (this results from the infinite divisibility property of the negative binomial).

$$Q = n(n + 2) \sum_{l=1}^L \frac{\hat{\rho}_l^2}{n - l}$$

with n being the sample size, $\hat{\rho}_l$ being the sample auto correlation coefficient at lag l , and L the total number of lags being tested. For a significance level α , the hypothesis of randomness is rejected if:

$$Q > \chi_{1-\alpha, L}^2$$

with $\chi_{1-\alpha, L}^2$ being the α -quantile of the chi-square distribution with L degrees of freedom.

High p -values are desirable, and values above 0.05 signify that we cannot reject the null hypothesis of independence of demand over time.

Appendix C. Analytical Model of Demand Recording

In this section, the goal is to derive the conditional probability distribution of the inventory on the shelf (or the discrepancy, since the book inventory is known), conditioned on operationally observable variables (sales and/or time since the last count). This may yield to improved counting and ordering heuristics. In this process, we verify that if the total demand is negative binomial, and recording occurs according to a Bernoulli process, then the recorded demand is negative binomial as well.

Notations and Assumptions

- D **daily** demand distribution
- p probability of each transaction being recorded
- D_R **daily** recorded demand distribution
- D_U **daily** unrecorded demand distribution
- τ time since the last count event, in days
- X **accumulated** total demand distribution since the last count event
- Y **accumulated** discrepancy, defined as the difference between the book inventory and the physical inventory on hand
- Z **accumulated** recorded demand distribution since the last count event

We make the following assumptions:

- The daily demand distribution follows a negative binomial distribution:
- $D \sim \text{NegBin}(\theta, u)$, i.i.d.

- We use the following parameterization of the Negative Binomial distribution $NegBin(\theta, u)$ with probability mass function f :

$$f_{Negbin}(failures = n, trials = \theta, success\ probability = u) = \binom{n + \theta - 1}{n} u^\theta (1-u)^n = \frac{\Gamma(n + \theta)}{n! \Gamma(\theta)} u^\theta (1-u)^n$$

- All demand transactions are unit-sized.
- The probability of a demand transaction being recorded is $0 < p < 1$ and follows a Bernoulli distribution. It is independent of demand.
- No product returns can take place.
- All counts are perfect: they reflect the “true” physical inventory, and therefore set $Y = 0$.
- The time since the last count is independent from the realized demand and the probability p , and is equal to τ

Results

From assumptions 1-5, we can derive the following results:

$$D_R \sim NegBin(\theta, x), \text{ i.i.d., with } x = \frac{u}{p+u-pu}$$

$$D_U \sim NegBin(\theta, y), \text{ i.i.d., with } y = \frac{u}{1-p+pu}$$

Additionally, using assumption 6:

$$X \sim NegBin(\theta\tau, u), \text{ i.i.d.}$$

$$Y \sim NegBin(\theta\tau, y), \text{ i.i.d.}$$

$$Z \sim NegBin(\theta\tau, x), \text{ i.i.d.}$$

$$Y | \tau, Z \sim NegBin(Z + \theta\tau, \frac{u}{x}), \text{ with } \frac{u}{x} = p + u - pu = 1 - (1-u)(1-p) = \frac{p}{1-x+xp}$$

If we assume that the original demand distribution is negative binomial, then the recorded and unrecorded demand also follows a negative binomial distribution with known assumptions. If the negative binomial distribution offers a good fit for the recorded demand distribution, it is plausible that the original demand distribution was also negative binomial.

The main benefit of this model is that it allows to infer the probability distribution of Y , the accumulated inventory error, based on the estimate of p , the estimates of the demand distribution parameters θ and x , the observed cumulative sales since the last count event Z , and the time since the last count event.

Derivation of Proposition 1

$$\begin{aligned}
 P[D_R = k] &= \sum_{n=0}^{+\infty} P[D_R = k \mid D = n] P[D = n] \\
 &= \sum_{n=k}^{+\infty} P[D_R = k \mid D = n] P[D = n] \\
 &= \sum_{n=k}^{+\infty} \left[\binom{n}{k} p^k (1-p)^{n-k} \right] \left[\frac{\Gamma(n+\theta)}{n! \Gamma(\theta)} u^\theta (1-u)^n \right]
 \end{aligned}$$

With a change of variable $i = n - k$, we have:

$$\begin{aligned}
 P[D_R = k] &= \sum_{i=0}^{+\infty} \left[\frac{1}{k! i!} p^k (1-p)^i \right] \left[\frac{\Gamma(i+k+\theta)}{\Gamma(\theta)} u^\theta (1-u)^{i+k} \right] \\
 &= \frac{p^k u^\theta (1-u)^k}{k! \Gamma(\theta)} \sum_{n=k}^{+\infty} \frac{\Gamma(i+k+\theta)}{i!} [(1-p)(1-u)]^i
 \end{aligned}$$

We apply the following changes of variables: $\theta' = k + \theta$, $1 - b = (1 - p)(1 - u)$,
 $b = 1 - (1 - p)(1 - u) = p + u - pu$

$$\begin{aligned}
 P[D_R = k] &= \frac{p^k u^\theta (1-u)^k}{k! \Gamma(\theta)} \frac{\Gamma(\theta')}{b^{\theta'}} \sum_{i=0}^{+\infty} \frac{\Gamma(i+\theta')}{i! \Gamma(\theta')} b^{\theta'} [1-b]^i \\
 &= \frac{p^k u^\theta (1-u)^k}{k! \Gamma(\theta)} \frac{\Gamma(\theta')}{b^{\theta'}} \\
 &= \frac{\Gamma(k+\theta)}{k! \Gamma(\theta)} \frac{p^k u^\theta (1-u)^k}{[p+u-pu]^{k+\theta}} \\
 &= \frac{\Gamma(k+\theta)}{k! \Gamma(\theta)} \left(\frac{u}{p+u-pu} \right)^\theta \left(\frac{p(1-u)}{p+u-pu} \right)^k
 \end{aligned}$$

With $x = \frac{u}{p+u-pu}$:

$$P[D_R = k] = \frac{\Gamma(k + \theta)}{k! \Gamma(\theta)} x^\theta [1 - x]^k$$

which is the p.d.f. of the $NegBin(\theta, x)$ distribution.

Derivation of other propositions

Using $p' = 1 - p$ and the result above, we derive Proposition 2 : $D_v \sim NegBin(\theta, y)$, i.i.d., with

$$y = \frac{u}{1-p+pu}$$

u and p can be expressed algebraically from x and y , and the proof of Propositions 3 and 4 is obtained by substituting the expressions for x and y .

Propositions 3-5 follow from the infinite divisibility property of the negative binomial distribution⁶⁵.

Proposition 6 requires further derivation. Using Bayes' theorem:

$$\begin{aligned} P[Y = s \mid \tau, Z = j] &= P[X = s + j \mid \tau, Z = j] \\ &= \frac{P[Z = j \mid \tau, X = s + j] P[X = s + j \mid \tau]}{\sum_{m=0}^{+\infty} P[Z = j \mid \tau, X = m] P[X = m \mid \tau]} \end{aligned}$$

We can calculate the numerator and the summed term in the denominator of this expression:

⁶⁵ A formal derivation can be obtained using Moment Generating Functions and is in Appendix A (to be added).

$$\begin{aligned}
P[Z = j | X = m] P[X = m | \tau] &= \left[\binom{m}{j} p^j (1-p)^{m-j} \right] \left[\frac{\Gamma(m + \theta\tau)}{m! \Gamma(\theta\tau)} u^{\theta\tau} (1-u)^m \right] \\
&= \frac{m!}{j!(m-j)!} \frac{\Gamma(m + \theta\tau)}{m! \Gamma(\theta\tau)} u^{\theta\tau} p^j (1-p)^{m-j} (1-u)^m \\
&= \frac{\Gamma(j + \theta\tau)}{j! \Gamma(\theta\tau)} \left[\frac{\Gamma(m - j + j + \theta\tau)}{(m-j)! \Gamma(j + \theta\tau)} u^{\theta\tau} p^j (1-p)^{m-j} (1-u)^m \right] \\
&= \frac{\Gamma(j + \theta\tau)}{j! \Gamma(\theta\tau)} \frac{(1-u)^j}{p^{\theta\tau} u^j} \left[\frac{\Gamma(m - j + j + \theta\tau)}{(m-j)! \Gamma(j + \theta\tau)} (pu)^{j+\theta\tau} [(1-p)(1-u)]^{m-j} \right] \\
&= \frac{\Gamma(j + \theta\tau)}{j! \Gamma(\theta\tau)} \frac{(1-u)^j}{p^{\theta\tau} u^j} \frac{(pu)^{j+\theta\tau}}{(1 - (1-p)(1-u))^{j+\theta\tau}} \\
&= \left[\frac{\Gamma(m - j + j + \theta\tau)}{(m-j)! \Gamma(j + \theta\tau)} [1 - (1-p)(1-u)]^{j+\theta\tau} [(1-p)(1-u)]^{m-j} \right] \\
&= \frac{\Gamma(j + \theta\tau)}{j! \Gamma(\theta\tau)} \frac{(1-u)^j}{p^{\theta\tau} u^j} \frac{(pu)^{j+\theta\tau}}{(1 - (1-p)(1-u))^{j+\theta\tau}} \\
&= f_{Negbin}(m - j, j + \theta\tau, [1 - (1-p)(1-u)])
\end{aligned}$$

Because it does not depend on m , the factor to the left of the expression is eliminated when taking the ratio:

$$P[Y = s | \tau, Z = j] = f_{Negbin}(s, j + \theta\tau, [1 - (1-p)(1-u)])$$

Extension for uneven demand rates across time

We look at the original demand distribution D , but since the distribution of recorded demand is also negative binomial, the results carry over to the distribution of recorded demand D_r . We make the following additional assumptions:

The day is divided in n time buckets of equal length.

Demand in each time bucket D_i is independent from demand in all other time buckets.

We relax the assumption that demand in all time buckets is identically distributed, and only require that:

$$D_i \sim \text{NegBin}(\theta\alpha_i, u)$$

Where $\alpha_i = \frac{E[D_i]}{E[D]}$, such that $\sum_{i=1}^{i=n} \alpha_i = 1$

Under those assumptions, we have:

$$\sum_{i=1}^{i=n} D_i \sim \text{NegBin}(\theta, u) \sim D$$

Let $S = \sum_{i=1}^{i=n} D_i$, and let $G_s(z), G_{D_i}(z)$ be the probability generating function of S and D_i .

Because the D_i are independent, we have:

$$G_s(z) = \prod_{i=1}^{i=n} G_{D_i}(z)$$

Substituting the probability generating function of the negative binomial, we have:

$$G_s(z) = \prod_{i=1}^{i=n} \left[\frac{u}{1-(1-u)z} \right]^{\theta\alpha_i} = \left[\frac{u}{1-(1-u)z} \right]^{\left(\sum_{i=1}^{i=n} \theta\alpha_i \right)} = \left[\frac{u}{1-(1-u)z} \right]^{\theta}$$

By identification of the probability generating function of the negative binomial, we have the desired result.

Appendix D. Discrete-Event Simulation Model Code

The following code is in the Python language, and uses the SimPy library for discrete-event simulation and numpy for random variable generation.

```
#from SimPy.SimPlot import *
from SimPy.Simulation import *
import sys, csv, time, math, random, numpy.random, gc
import MySQLdb

class StaSKU:
    def __init__(self,L,seed,hashCode,trial):

        self.hashCode = hashCode
        self.seed = seed
        self.trial = trial

        # Activate processes
        sd = SourceDemand()
        activate(sd, sd.runSD(L, self))

        # Initialize the count process
        # There are different possible values for the IntCt Field
        # 0: [a number] Periodic counts. as a convention, enter 100000
        # as a way to guarantee no periodic counts at all
        # 1: [a distribution, such as "exp(4.0)" for exponentially
        # distributed inter-count times with mean 4
        # or "norm(mean,sd)" e.g. "norm(4;0.1)" Note the use of the
        # semicolon to avoid problems with .csv files
        # 2: [ a .csv file, in order to load historical data.
```

```

    # function(A,B), where function = norm(mean, std, trunc-at-
2*mean=True or False) ; exp(mean), unif(A,B)

    # The line below is removed we may want to count on the book
view inventory being below the reorder point
    # AND periodically as well
    # if eval(L['CountLow']) == 0

    if L['Inter-Count time'].endswith(".csv"):
        # Historical inter count times are loaded from a .csv file
        # If the sum of the inter-count times provided is below
the length of the simulation run,
        # the system will loop over the historical file
        self.ctType = 2
        f3 = open(L['Inter-Count time'], "rU")
        self.historCtData = []
        for line in csv.reader(f3.readlines()[1:]):
            self.historCtData.append(line)
        f3.close()

    elif L['Inter-Count time'].endswith(""):
        # We use a distribution in this case.
        # The different types of distributions recognized are defined
below
        self.ctType = 1

        # stream for the distribution of inter-count arrival times
        self.IntCtRandom = random.Random(self.seed)

        # Parse string to determine distribution time and parameters
        self.IntCtDistType, sep, ab = L['Inter-Count time'][:-
1].partition('(')

        self.IntCtA, sep, ab2 = ab.partition(';')
        self.IntCtB, sep, self.IntCtC = ab2.partition(';')

    elif eval(L['Inter-Count time']) > 0:
        self.ctType = 0

    else:
        print("error: IntCt field syntax error")
        sys.exit()

    # Finally, initialize the count process

```

```

sc = SourceCount()
activate(sc,sc.runSC(L,self),at=eval(L['Initial count time']))

ord = Order()
activate(ord,ord.runO(L, self))
rev = Review()
activate(rev,rev.runRev(L,self,ord))

# Define levels, monitors, lists
self.realInvLev = Level(initialBuffered=eval(L['InitialIOH']),
monitored=True) # Level indicates real
# inventory amount
self.monRealFull = Monitor(name='%s' % (L['Demand Error probability']),ylab = 'real inv. level') # monitor real inventory level
over entire simulation
self.monReal = self.realInvLev.bufferMon # monitor real inventory level over each statTimeRange
self.aveRealInv = [] # list of time-averaged real inv. level per statTimeRange

self.BVinv = eval(L['InitialIOH']) # book view inventory level
self.monBVFull = Monitor(name='%s' % (L['Demand Error probability']),ylab = 'book inv. level') # monitor book view inventory level
over entire simulation
self.monBV = Monitor() # monitor book view inventory level over each statTimeRange
self.aveBVInv = [] # list of time-averaged BV inv. level per statTimeRange

self.monRVSOFull = Monitor(ylab = 'RVSO') # monitor # of real stockouts over entire simulation
self.monRVSO = Monitor() # monitor # of real stockouts over each statTimeRange
self.sumRVSO = [] # list of sum of real stockouts per statTimeRange

self.monBVSOFull = Monitor(ylab = 'BVSO') # monitor # of recorded stockouts (book view) over entire simulation
self.monBVSO = Monitor() # monitor # of recorded stockouts (book view) over each statTimeRange
self.sumBVSO = [] # list of sum of book view stockouts per statTimeRange

```

```
self.monLostSaleFull = Monitor(ylab = 'LostSales') # monitor
amt of lost sales over entire simulation
self.monLostSale = Monitor() # monitor amt of lost sales over
entire each statTimeRange
self.sumLostSale = [] # list of sum of lost sales per stat-
TimeRange

self.monKilledFull = Monitor(ylab = '# Killed Orders') # moni-
tor # of killed orders over entire simulation
self.monKilled = Monitor() # monitor # of killed orders over
each statTimeRange
self.sumKilled = [] # list of sum of killed orders per stat-
TimeRange

self.monRVDemFull = Monitor(ylab = 'RVDemand') # monitor quan-
tity demanded every time bucket
# over entire simulation
self.monRVDem = Monitor() # monitor quantity demanded every
time bucket over each statTimeRange
self.sumRVDem = [] # list of sum of demand quantities per
statTimeRange

self.monRVSalesFull = Monitor(ylab = 'RVSales') # monitor
quantity removed from real inventory
# at each demand instance (sales) over entire simulation
self.monRVSales = Monitor()
self.sumRVSales = []

self.monBVSalesFull = Monitor(ylab = 'BVSales') # monitor
quantity recorded to have been
# removed from real inventory at each demand instance (sales)
over entire simulation
self.monBVSales = Monitor()
self.sumBVSales = []

# monitor number of times that the count transaction occurs
self.monNumCts = Monitor()
self.sumNumCts = []

self.monNumReplFull = Monitor(ylab = 'NumReplenTrans') # moni-
tor number of times that
# the replenishment transaction occurs over entire simulation
self.monNumRepl = Monitor()
self.sumNumRepl = []

self.BVavailFull = Monitor(name='BVavailfull, %s' % (L['Demand
```

```

Error probability'])) # monitor indicates non-zero real inventory (binary var)
    # over entire simulation
    self.BVavailMon = Monitor() # monitor indicates non-zero book view inventory (binary var),
    # reset every statTimeRange
    self.RVavailPerc = [] # list of % of time RV inventory is non-zero

    self.RVavailFull = Monitor(name='RVavailfull, %s' % (L['Demand Error probability'])) # monitor indicates non-zero book view inventory (binary var)
    # reset every statTimeRange
    self.RVavailMon = Monitor() # monitor indicates non-zero real inventory (binary var),
    # reset every statTimeRange
    self.BVavailPerc = [] # list of % of time BV inventory is non-zero

    self.BVnegMon= Monitor()
    self.BVPercNeg = []

    self.lastRVSO = [] # keeps track of when the most recent RVSO, that did not happen concurrently with a BVSO, occurred

    self.transQuant = Monitor() # monitor tracks the sales and replenishment quantities, reset every statTimeRange
    self.transQuantSum = [] # list of sums of transQuant every statTimeRange
    self.absDiscrep = Monitor() # monitor tracks the absolute value of the discrepancies between the
    # real and book view discovered during a Count, reset every statTimeRange
    self.absDiscrepSum = [] # list of sums of absDiscrep every statTimeRange

    self.demSinceCt = Monitor() # monitor tracks sales that have occurred since last Count event,
    # reset every time Count is called
    self.p_estDemMon = Monitor()
    self.p_estDiscrepMon = Monitor()
    self.p_estDemList = []
    self.p_estDiscrepList = []

    self.retSinceCt = Monitor()
    self.CountRetMon = Monitor()

```

```

self.replSinceCt = Monitor()
self.CountReplMon = Monitor()
self.RVRetMon = Monitor()
self.RVReplMon = Monitor()
self.BVRetMon = Monitor()

if eval(L['InitialIOH']) == 0:
    self.tSinceRVSO = 0
    self.tSinceBVSO = 0
else:
    self.tSinceRVSO = None
    self.tSinceBVSO = None
self.RVSOdurMon = Monitor()
self.RVSOdurList = []
self.BVSOdurMon = Monitor()
self.BVSOdurList = []

ts = timeStats() # Initialize process that resets short term
monitors every statTimeRange
activate(ts,ts.runTS(self))

self.activeOrders = [] # list of currently active orders

##          Define variables specific to Station SKU and save in dic-
tionary "StaSKUVars"
##          self.hashCode = hash(hash(str(time.time()))+hash(str(L))) #
assign hash code that
##          is unique to each Station SKU that is simulated. The hash
code is a function of the current time
##          and of the dictionary of inputs. Therefore, different runs
will not produce the same hash code.
##          This code will be used to identify and index the simulated
Station SKUs in the MySQL database.
self.runTime = str(time.ctime())
self.StaSKUVars = {'FacilitySta-
tion':L['Station'],'Minimum':L['Minimum'],'Maximum':L['Maximum'],\
'Item
Class':L['ItemClass'],'ItemID':L['ItemID'],'Run Time':self.runTime}
# This dictionary is used to store the variables that are spe-
cific to a Station SKU, but do not change with each transaction

# Run method that will write all parameters, global and Sta-
tion-SKU specific, to an output file
pw = paramsWriter()
pw.pwRun(L,self)

```

```

    # Initialize random streams for each source of uncertainty
    self.demRand = numpy.random.RandomState() # This object ini-
tializes the random stream for the quantity
    # demanded each time bucket.
    self.demRand.seed(self.seed)
    self.demRec = random.Random(self.seed) # This object initial-
izes the random
    # stream for the Demand Error (whether a Demand transaction is
recorded)
    self.replRand = random.Random(self.seed) # This object ini-
tializes the random
    # stream for the Replenishment error (whether Count is called
when an order is filled or not)
    self.returnTime = random.Random(self.seed) # This object ini-
tializes the random
    # stream for the time between a demand and return for each re-
turn transaction
    self.returnRec = random.Random(self.seed) # This object ini-
tializes the random
    # stream for the Return error (whether a Return is recorded or
not)
    self.returnProb = random.Random(self.seed) # This object ini-
tializes the random
    # stream for whether each unit of sales is returned
    self.CtOnLS = random.Random(self.seed)
    # stream for whether a count takes place whenever there is a
lost sale (i.e. unmet demand)

    # Save transactional data in lists stored in the dictionary
"transRec"
    self.transRec = { 'Hash-
Code':self.hashCode, 'Seed':self.seed, 'Time':[], 'Transaction
Type':[], 'Quantity':[],
                    'BV Expected Count':[], 'BV Actual
Count':[], 'BV End Count':[], 'RV Expected Count':[], \
                    'RV End
Count':[], 'Recorded':[], 'CountEventTrue':[], 'BV StockoutEvent':[], \
                    'RV StockoutE-
vent':[], 'BVSOLag':[], 'Trial':self.trial} # This dictionary of lists
is used to store all of the transactional data.

# This guaranteed that station SKU records got dumped to the

```

```

file periodically
#         Now that we only run one station sku trial at a time, there
is no point in using it
#         lw = logWriter()
#         activate(lw,lw.lwRun(self))

        self.startList =[]
        self.endList = []
        self.CountRetSum = []; self.CountReplSum = []; self.RVReplSum
= [];
        self.RVRetSum = []; self.BVRetSum = []

class paramsWriter:
    def pwRun(self,L, stasku):
        constVars-
Rec.writerow(("N", stasku.hashCode, stasku.StaSKUVars['FacilityStation'
],
stasku.StaSKUVars['ItemID'], stasku.StaSKUVars['Minimum'], stasku.StaSKU
Vars['Maximum'], \
                stasku.runTime, stasku.StaSKUVars['Item
Class'], L['InitialIOH'], L['Demand Lambda'], \
L['a0'], L['a1'], L['a2'], L['a3'], L['a4'], L['a5'], L['a6'], L['a7'], \
                L['a8'], L['a9'], L['a10'], L['a11'], L['Demand Er-
ror probability'], \
                L['Initial        count        time'], L['Inter-Count
time'], "N", \
                L['kill time'], L['Replenish Error probabil-
ity'], \
                globalDict['Run        length'], globalDict['Initial
time bucket'], L['NegBinProb'], \
                globalDict['Initial        day        of
week'], "N", "N", "N", "N", "N", "N", \
"N", L['LT1'], L['LT2'], L['LT3'], L['LT4'], L['LT5'], L['LT6'], L['LT7'], \
L['ReturnProb'], L['ReturnMeanTime'], stasku.trial, stasku.seed, \
L['R0'], L['R1'], L['R2'], L['R3'], L['R4'], L['R5'], L['R6'], \
                globalDict['statTimeRange'], L['Return        Error
probability'], L['CountLow'], L['probCtOnLS'], globalDict['NumTrials']))

class logWriter(Process):
    def lwRun(self, stasku):

```



```

while True:
    yield hold, self, 49
    self.twrite(stasku)
    self.swrite(stasku)
def twrite(self, stasku):
    for m in range(len(stasku.transRec['Transaction Type'])):
        if eval(globalDict['LagRec']) == 1:
            transLog.writerow((stasku.transRec['HashCode'],\
stasku.transRec['Time'][m],stasku.transRec['Transaction Type'][m],\
stasku.transRec['Quantity'][m],stasku.transRec['BV Expected
Count'][m],\
                                stasku.transRec['BV Actual
Count'][m],stasku.transRec['BV End Count'][m],\
                                stasku.transRec['RV Expected
Count'][m],stasku.transRec['RV End Count'][m],\
stasku.transRec['Recorded'][m],stasku.transRec['CountEventTrue'][m],\
                                stasku.transRec['RV StockoutE-
vent'][m],stasku.transRec['BV StockoutEvent'][m],\
"\N",stasku.transRec['BVSOLag'][m],stasku.trial))
        else:
            transLog.writerow((stasku.transRec['HashCode'],\
stasku.transRec['Time'][m],stasku.transRec['Transaction Type'][m],\
stasku.transRec['Quantity'][m],stasku.transRec['BV Expected
Count'][m],\
                                stasku.transRec['BV Actual
Count'][m],stasku.transRec['BV End Count'][m],\
                                stasku.transRec['RV Expected
Count'][m],stasku.transRec['RV End Count'][m],\
stasku.transRec['Recorded'][m],stasku.transRec['CountEventTrue'][m],\
                                stasku.transRec['RV StockoutE-
vent'][m],stasku.transRec['BV StockoutEvent'][m],\
                                "\N", "\N",stasku.trial))
            # Reinitialize all lists after written to csv file
            stasku.transRec = { 'Hash-
Code':stasku.hashCode, 'Seed':stasku.seed, 'Time':[], 'Transaction
Type':[], 'Quantity':[],
                                'BV Expected Count':[], 'BV Actual Count':[], 'BV End
Count':[], 'RV Expected Count':[],\
                                'RV End Count':[], 'Recorded':[], 'CountEventTrue':[], 'BV

```

```

StockoutEvent': [], \
    'RV StockoutEvent': [], 'BVSOLag': [], 'Trial': stasku.trial}
def swrite(self, stasku):
    for i in range(len(stasku.sumBVSales)):
        # Composite metrics:
        if stasku.sumRVSales[i] != 0:
            RecProb = float(stasku.sumBVSales[i]/stasku.sumRVSales[i])
            RVSORate = float(stasku.sumRVSO[i]/stasku.sumRVSales[i])
        else:
            RecProb = "\N"; RVSORate = "\N"
        if stasku.sumBVSales[i] != 0:
            BVSORate = float(stasku.sumBVSO[i]/stasku.sumBVSales[i])
        else:
            BVSORate = "\N"
        if stasku.sumRVDem[i] != 0:
            FillRate = float(stasku.sumLostSale[i]/stasku.sumRVDem[i])
        else:
            FillRate = "\N"
        stats = Rec.writerow((stasku.hashCode, stasku.startList[i], stasku.endList[i], stasku.sumRVDem[i], stasku.sumLostSale[i], \
            stasku.sumRVSales[i], stasku.sumBVSales[i], stasku.sumRVSO[i], stasku.sumBVSO[i], stasku.sumNumCts[i], \
            stasku.sumNumRepl[i], stasku.RVavailPerc[i], stasku.BVavailPerc[i], stasku.aveRealInv[i], stasku.aveBVInv[i], \
            stasku.sumKilled[i], RecProb, FillRate, RVSORate, BVSORate, "\N", stasku.transQuantSum[i], stasku.absDiscrepSum[i], \
            stasku.p_estDemList[i], stasku.p_estDiscrepList[i], stasku.BVPercNeg[i], stasku.CountRetSum[i], \
            stasku.CountReplSum[i], stasku.RVReplSum[i], stasku.RVRetSum[i], stasku.BVRetSum[i], \
            stasku.RVSOdurList[i], stasku.BVSOdurList[i], stasku.trial))
            stasku.startList = []; stasku.endList = []; stasku.sumRVDem = []; stasku.sumLostSale = [];
            stasku.sumRVSales = []; stasku.sumBVSales = []; stasku.sumRVSO = []; stasku.sumBVSO = []; stasku.sumNumCts = [];

```

```

        stasku.sumNumRepl      =      [];      stasku.RVavailPerc      =      [];
stasku.BVavailPerc = []; stasku.aveRealInv = []; stasku.aveBVInv = [];
        stasku.sumKilled      =      [];      stasku.transQuantSum      =      [];
stasku.absDiscrepSum = [];
        stasku.p_estDemList    =      [];      stasku.p_estDiscrepList    =      [];
stasku.BVPercNeg = [];
        stasku.CountRetSum     =      [];      stasku.CountReplSum     =      [];
stasku.RVReplSum = [];
        stasku.RVRetSum = []; stasku.BVRetSum = []; stasku.RVSOdurList
= []; stasku.BVSOdurList = []

```

```

class timeStats(Process):
    def runTS(self, stasku):
        while True:
            # For monitors of continuous variables (real and BV inven-
            tory and availability),
            # collect data points at the beginning and end of each
            statistical analysis period
            # (called "statTimeRange")
            stasku.monReal.observe(stasku.realInvLev.amount)
            stasku.monBV.observe(stasku.BVinv)
            if stasku.BVinv < 0:
                stasku.BVnegMon.observe(1)
            else:
                stasku.BVnegMon.observe(0)
            stasku.startList.append(now())
            yield hold, self, eval(globalDict['statTimeRange'])
            stasku.endList.append(now())
            stasku.RVavailPerc.append(stasku.RVavailMon.timeAverage())
            stasku.BVavailPerc.append(stasku.BVavailMon.timeAverage())
            stasku.monReal.observe(stasku.realInvLev.amount)
            stasku.aveRealInv.append(stasku.monReal.timeAverage())
            stasku.monBV.observe(stasku.BVinv)
            stasku.aveBVInv.append(stasku.monBV.timeAverage())
            stasku.BVPercNeg.append(stasku.BVnegMon.timeAverage())
            if len(stasku.RVSOdurMon) > 0 :
                stasku.RVSOdurList.append(stasku.RVSOdurMon.mean())
            else:
                stasku.RVSOdurList.append("\N")
            if len(stasku.BVSOdurMon) > 0:
                stasku.BVSOdurList.append(stasku.BVSOdurMon.mean())
            else:
                stasku.BVSOdurList.append("\N")
            # For monitors of flow variables (real and BV stockout
            events, demand quantities,

```

```
    # lost sales, number count and replenishment transactions,
and # of killed orders),
    # collect sum of quantities over each statTimeRange
    stasku.sumRVSO.append(stasku.monRVSO.total())
    stasku.sumBVSO.append(stasku.monBVSO.total())
    stasku.sumLostSale.append(stasku.monLostSale.total())
    stasku.sumRVDem.append(stasku.monRVDem.total())
    stasku.sumRVSales.append(stasku.monRVSales.total())
    stasku.sumBVSales.append(stasku.monBVSales.total())
    stasku.sumNumCts.append(stasku.monNumCts.total())
    stasku.sumNumRepl.append(stasku.monNumRepl.total())
    stasku.sumKilled.append(stasku.monKilled.total())
    # Performance metrics:
    stasku.transQuantSum.append(stasku.transQuant.total())
    stasku.absDiscrepSum.append(stasku.absDiscrep.total())
    stasku.p_estDemList.append(stasku.p_estDemMon.total())

stasku.p_estDiscrepList.append(stasku.p_estDiscrepMon.total())
    stasku.CountRetSum.append(stasku.CountRetMon.total())
    stasku.CountReplSum.append(stasku.CountReplMon.total())
    stasku.RVReplSum.append(stasku.RVReplMon.total())
    stasku.RVRetSum.append(stasku.RVRetMon.total())
    stasku.BVRetSum.append(stasku.BVRetMon.total())

    # Reset each monitor each statTimeRange
    stasku.RVavailMon.reset()
    stasku.BVavailMon.reset()
    stasku.monReal.reset()
    stasku.monBV.reset()
    stasku.monRVSO.reset()
    stasku.monBVSO.reset()
    stasku.monLostSale.reset()
    stasku.monRVDem.reset()
    stasku.monRVSales.reset()
    stasku.monBVSales.reset()
    stasku.monNumCts.reset()
    stasku.monNumRepl.reset()
    stasku.monKilled.reset()
    stasku.transQuant.reset()
    stasku.absDiscrep.reset()
    stasku.p_estDemMon.reset()
    stasku.p_estDiscrepMon.reset()
    stasku.BVnegMon.reset()
    stasku.CountRetMon.reset()
    stasku.CountReplMon.reset()
    stasku.RVReplMon.reset()
```

```

        stasku.RVRetMon.reset()
        stasku.BVRetMon.reset()
        stasku.RVSOdurMon.reset()
        stasku.BVSOdurMon.reset()

class SourceDemand(Process):
    def runSD(self, L, stasku):
        # Generate instance of Demand every time bucket
        while True:
            d = Demand()
            activate(d, d.runD(L, stasku))
            yield hold, self, tb

class Demand(Process):
    def runD(self, L, stasku):

        # Demand Process Inputs
        # Assuming negative binomial distribution for # of StaSKUs de-
        # manded every time bucket
        # where the probability is constant but the number of trials
        # changes as per the fraction of the
        # daily demand for each time bucket
        negBinP = eval(L['NegBinProb'])
        sizeParam = eval(L['Demand Lambda'])*negBinP/(1-negBinP) #
        mean # of StaSKUs demanded every day
        a = [eval(L['a0']),eval(L['a1']),eval(L['a2']),eval(L['a3']),\
            eval(L['a4']),eval(L['a5']),eval(L['a6']),eval(L['a7']),\
            eval(L['a8']),eval(L['a9']),eval(L['a10']),eval(L['a11'])] # fraction of

        # daily demand that occurs during each bucket (position in
        # vector indicates time bucket)
        if a[t] == 0: # negative binomial distribution does not apply
            when the number of trials is 0
                Dt = 0
            else:
                Dt =
                =
        stasku.demRand.poisson(stasku.demRand.gamma(sizeParam*a[t], (1-
        negBinP)/negBinP)) # quantity demanded each time bucket
        stasku.monRVDem.observe(Dt)
        if eval(globalDict['Plot']) == 1:
            stasku.monRVDemFull.observe(Dt)
        p = eval(L['Demand Error probability']) # probability that de-
        mand is not recorded, [0,1]
        e = stasku.demRec.random()

```

```

    # Demand Process
    qd = min(Dt, stasku.realInvLev.amount) # sales, quantity re-
moved from the machine
    stasku.monRVSales.observe(qd)
    if eval(globalDict['Plot']) == 1:
        stasku.monRVSalesFull.observe(qd)
    BVExpCt = stasku.BVinv # book view expected count (prior to
demand transaction)
    BVActCt = stasku.BVinv # book view actual count (same as ex-
pected for v2)
    RVEExpCt = stasku.realInvLev.amount # real view expected count
(prior to demand transaction)
    RVEndCt = RVEExpCt - qd # real view end count
    yield get, self, stasku.realInvLev, qd # remove sales from the
real inventory level
    if eval(globalDict['Plot']) == 1:
        stasku.monRealFull.observe(stasku.realInvLev.amount)
    if stasku.realInvLev.amount == 0: # real stockout event occurs
        RVSOevent = 1
    else: # real stockout event does NOT occur
        RVSOevent = 0
    if e >= p: # record the demand transaction
        stasku.monBVSales.observe(qd)
        stasku.demSinceCt.observe(qd) # record demand since last
count event
        stasku.transQuant.observe(qd) # record transaction quan-
tity

    if eval(globalDict['Plot']) == 1:
        stasku.monBVSalesFull.observe(qd)
        recorded = 1
    BVEndCt = BVExpCt - qd # book view end count
    stasku.BVinv -= qd # remove sales from book view inventory
level

    stasku.monBV.observe(stasku.BVinv)
    if eval(globalDict['Plot']) == 1:
        stasku.monBVFull.observe(stasku.BVinv)
    if BVEndCt <= 0: # book view stockout event occurs
        BVSOevent = 1
    else: # book view stockout event does not occur
        BVSOevent = 0
    else: # demand transaction is NOT recorded
        recorded = 0
        BVEndCt = BVExpCt # book view end count
        BVSOevent = 0
    if BVEndCt < 0:
        stasku.BVnegMon.observe(1)

```

```

else:
    stasku.BVnegMon.observe(0)
if Dt <= RVExpCt:
    ls = 0 # if real inventory can satisfy demand, there is 0
lost sales
else:
    ls = Dt - qd # if real inventory cannot satisfy demand,
lost sales is the difference
# between the quantity demanded and the sales supplied
from the inventory
stasku.monLostSale.observe(ls)
if eval(globalDict['Plot']) == 1:
    stasku.monLostSaleFull.observe(ls)
# If book view stockout begins, record start time
if BVExpCt > 0 and BVEndCt <= 0:
    stasku.tSinceBVSO = now()
# If real view stockout begins, record start time
if RVExpCt > 0 and RVEndCt == 0:
    stasku.tSinceRVSO = now()

# For each lost sales event, call a Count event with some
probability
if ls > 0:
    CET = 0 # Initialize CountEventTrue variable to 0
    c_rand = stasku.CtOnLS.random()
    BVActCtLS = BVExpCt
    BVEndCtLS = BVExpCt
    if c_rand <= eval(L['probCtOnLS']):
        c = Count()
        c.runC(L, stasku)
        CET = 1
        BVActCtLS = RVExpCt
        BVEndCtLS = RVExpCt

# Record transactional data for Demand and/or Lost Sales
transactions
if qd > 0: # if sales > 0, record transaction in log
    BVSOLag = "\N"
    if RVSOevent == 1:
        stasku.lastRVSO.append(now())
        if BVSOevent == 1:
            BVSOLag = 0
    if eval(globalDict['LagRec']) == 1:
        stasku.transRec['BVSOLag'].append(BVSOLag)
    stasku.monRVSO.observe(RVSOevent)
    stasku.monBVSO.observe(BVSOevent)

```

```

if eval(globalDict['Plot']) == 1:
    stasku.monRVSOFull.observe(RVSOevent)
    stasku.monBVSOFull.observe(BVSOevent)
stasku.transRec['Time'].append(now())
stasku.transRec['Transaction Type'].append(1)
stasku.transRec['Quantity'].append(qd)
stasku.transRec['BV Expected Count'].append(BVExpCt)
stasku.transRec['BV Actual Count'].append(BVActCt)
stasku.transRec['BV End Count'].append(BVEndCt)
stasku.transRec['RV Expected Count'].append(RVExpCt)
stasku.transRec['RV End Count'].append(RVEndCt)
stasku.transRec['Recorded'].append(recorded)
stasku.transRec['CountEventTrue'].append(0)
stasku.transRec['BV StockoutEvent'].append(BVSOevent)
stasku.transRec['RV StockoutEvent'].append(RVSOevent)
if ls > 0: # if lost sales > 0, record transaction in log
    stasku.transRec['Time'].append(now())
    stasku.transRec['Transaction Type'].append(21)
    stasku.transRec['Quantity'].append(ls)
    # no exchange of items in lost sales transaction, no
change made to inventory
    # amounts below
    stasku.transRec['BV Expected Count'].append(BVExpCt)
    stasku.transRec['BV Actual Count'].append(BVActCtLS)
    stasku.transRec['BV End Count'].append(BVEndCtLS)
    stasku.transRec['RV Expected Count'].append(RVExpCt)
    stasku.transRec['RV End Count'].append(RVEndCt)
    stasku.transRec['Recorded'].append(0) # lost sales is not
recorded in transaction log from machine
    stasku.transRec['CountEventTrue'].append(CET)
    stasku.transRec['BV StockoutEvent'].append(0) # this quan-
tity has no meaning for a lost sales transaction
    stasku.transRec['RV StockoutEvent'].append(0) # this quan-
tity has no meaning for a lost sales transaction
    stasku.transRec['BVSOLag'].append("\N")
    # Determine how many items, of the sales quantity, will be re-
turned
    if qd != 0:
        qRet = 0 # initialize the number of items to be returned
        for i in range(qd):
            # Each unit of sales has probability of being re-
turned, independent of whether
            # the other units of sales are returned
            ret = stasku.returnProb.random()
            if ret < eval(L['ReturnProb']):
                qRet += 1

```



```

        if qRet > 0: # if non-zero number of items are returned,
call Return process
            R = ReturnP()
            activate(R,R.runRet(L, stasku, qRet))

# Availability check
if stasku.realInvLev.amount == 0:
    stasku.RVavailMon.observe(0)
    if eval(globalDict['Plot']) == 1:
        stasku.RVavailFull.observe(0)
else:
    stasku.RVavailMon.observe(1)
    if eval(globalDict['Plot']) == 1:
        stasku.RVavailFull.observe(1)
if stasku.BVinv <= 0:
    stasku.BVavailMon.observe(0)
    if eval(globalDict['Plot']) == 1:
        stasku.BVavailFull.observe(0)
else:
    stasku.BVavailMon.observe(1)
    if eval(globalDict['Plot']) == 1:
        stasku.BVavailFull.observe(1)

class ReturnP(Process):
    def runRet(self,L, stasku, r):
        wT = stasku.returnTime.expovariate(lambd =
float(1.0/eval(L['ReturnMeanTime'])))
        yield hold, self, wT
        RVExpCt = stasku.realInvLev.amount # real view expected count
(prior to Return transaction)
        BVExpCt = stasku.BVinv # book view expected count (prior to
Return transaction)
        BVActCt = stasku.BVinv # book view actual count (same as ex-
pected for v2)
        RVEndCt = RVExpCt + r # real view end count

        yield put, self, stasku.realInvLev, r
        stasku.RVRetMon.observe(r)
        if eval(globalDict['Plot']) == 1:
            stasku.monRealFull.observe(stasku.realInvLev.amount)
        rE = stasku.returnRec.random() # Determine if Return transac-
tion will be recorded
        if rE >= eval(L['Return Error probability']):
            stasku.transQuant.observe(r)
            stasku.retSinceCt.observe(r)
            stasku.BVRetMon.observe(r)

```

```

    BVEndCt = BVExpCt + r # book view end count
    stasku.BVinv += r
    stasku.monBV.observe(stasku.BVinv)
    retRec = 1
    if eval(globalDict['Plot']) == 1:
        stasku.monBVFull.observe(stasku.BVinv)
else:
    BVEndCt = BVExpCt # book view end count
    retRec = 0
# If return transaction ends BV stockout, record BVSO duration
if BVExpCt <= 0 and BVEndCt > 0:
    stasku.BVSOdurMon.observe(now() - stasku.tSinceBVSO)
    stasku.tSinceBVSO = None
# If return transaction ends RV stockout, record RVSO duration
if RVExpCt == 0 and RVEndCt > 0:
    stasku.RVSOdurMon.observe(now() - stasku.tSinceRVSO)
    stasku.tSinceRVSO = None

# Record transactional data for Return transactions
stasku.transRec['Time'].append(now())
stasku.transRec['Transaction Type'].append(3) # "Transaction
Type" = 3 indicates Return transaction
stasku.transRec['Quantity'].append(r)
stasku.transRec['BV Expected Count'].append(BVExpCt)
stasku.transRec['BV Actual Count'].append(BVActCt)
stasku.transRec['BV End Count'].append(BVEndCt)
stasku.transRec['RV Expected Count'].append(RVExpCt)
stasku.transRec['RV End Count'].append(RVEndCt)
stasku.transRec['Recorded'].append(retRec)
stasku.transRec['CountEventTrue'].append(0) # For v2, assume
that a Count never accompanies a Return
stasku.transRec['RV StockoutEvent'].append(0) # a stockout can
never occur with a Return transaction
stasku.transRec['BV StockoutEvent'].append(0) # a stockout can
never occur with a Return transaction
stasku.transRec['BVSOLag'].append("\N")

class SourceCount(Process):
    def runSC(self, L, stasku):
        counter = 0
        while True:
            BVExpCt = stasku.BVinv # book view expected count (prior
to Count transaction)
            BVErr = stasku.realInvLev.amount - stasku.BVinv # dis-
crepancy between real and book view inventory levels

```

```

        BVActCt = stasku.BVinv + BVerror # book view actual count,
fix discrepancy
        BVEndCt = BVActCt # book view end count
        RVExpCt = stasku.realInvLev.amount # real view expected
count (prior to Count transaction)
        RVEndCt = RVExpCt # real view end count
        # record book view stockout only if the book view level is
changed to 0 as a result of this count event.
        # the book view level will only be changed to 0 if it is
currently not at 0, and the real view currently
        # is at a level of 0
        BVSOLag = "\N"
        if stasku.realInvLev.amount == 0 &
stasku.realInvLev.amount != stasku.BVinv:
            BVSOLag = 1
            if eval(globalDict['LagRec']) == 1:
                BVSOLag = now()-stasku.lastRVSO.pop()
                stasku.monBVSOLag.observe(BVSOLag)
            if eval(globalDict['Plot']) == 1:
                stasku.monBVSOLagFull.observe(BVSOLag)
        else: # conditions not met, do NOT record a book view
stockout
            BVSOLag = 0
        c = Count()
        c.runC(L, stasku)

        # Record transactional data for Count transactions (only
when called by SourceCount)
        stasku.transRec['Time'].append(now())
        stasku.transRec['Transaction Type'].append(4) # "Transac-
tion Type" = 4 indicates Count transaction
        stasku.transRec['Quantity'].append(0) # quantity has no
meaning for Count transaction
        stasku.transRec['BV Expected Count'].append(BVExpCt)
        stasku.transRec['BV Actual Count'].append(BVActCt)
        stasku.transRec['BV End Count'].append(BVEndCt)
        stasku.transRec['RV Expected Count'].append(RVExpCt)
        stasku.transRec['RV End Count'].append(RVEndCt)
        stasku.transRec['Recorded'].append(1) # assume that Count
transaction is always recorded
        stasku.transRec['CountEventTrue'].append(1)
        stasku.transRec['RV StockoutEvent'].append(0)
        stasku.transRec['BV StockoutEvent'].append(BVSOLag)
        if eval(globalDict['LagRec']) == 1:
            stasku.transRec['BVSOLag'].append(BVSOLag)

```

```

        # Check if counts have fixed periodicity or just happen at
predetermined inter-count times.
        if stasku.ctType == 2:
            # If we have looped through all inter-count times,
then we go back to the beginning
            if counter == len(stasku.historCtData):
                counter = 0
                print("Run length is greater than sum of loaded
InterCount times --> looping through them.")
                timetohold = eval(stasku.historCtData[counter][0])
                counter += 1
            elif stasku.ctType == 1:
                # Parse string to determine distribution time and pa-
rameters

                if stasku.IntCtDistType == "norm":
                    truncatecondition = True

                    while truncatecondition:
                        # We truncate the normal by resampling when-
ever it is value is below zero or above 2 times its mean, in order to
control for changes in the mean
                            timetohold =
stasku.IntCtRandom.normalvariate(float(eval(stasku.IntCtA)),float(eval
(stasku.IntCtB)))
                            if stasku.IntCtC == "True":
                                truncatecondition = (timetohold < 0) |
(timetohold > 2*float(eval(stasku.IntCtA)))
                            else:
                                truncatecondition = (timetohold < 0)
                    elif stasku.IntCtDistType == "exp":
                        timetohold =
stasku.IntCtRandom.expovariate(float(1.0/eval(stasku.IntCtA)))
                    elif stasku.IntCtDistType == "unif":
                        timetohold =
stasku.IntCtRandom.uniform(max(float(eval(stasku.IntCtA)),0.0),max(flo
at(eval(stasku.IntCtB)),0.0))
                            print("IntCt Parameter in uniform dist takes on
negative value, adjusted to zero")
                    else:
                        print("Error in parsing the function expression
for IntCt time")
                            print(stasku.transRec)
                            timetohold = -1
                else:
                    timetohold = eval(L['Inter-Count time']) # hold for
fixed time between count events

```

```

        #print timetohold
        # Yield for timetohold
        yield hold, self, timetohold

class SourceCountLow:
    def runSCL(self, L, stasku):
        BVExpCt = stasku.BVinv # book view expected count (prior
to Count transaction)
        BVError = stasku.realInvLev.amount - stasku.BVinv # dis-
crepancy between real and book view inventory levels
        BVActCt = stasku.BVinv + BVError # book view actual count,
fix discrepancy
        BVEndCt = BVActCt # book view end count
        RVExpCt = stasku.realInvLev.amount # real view expected
count (prior to Count transaction)
        RVEndCt = RVExpCt # real view end count
        # record book view stockout only if the book view level is
changed to 0 as a result of this count event.
        # the book view level will only be changed to 0 if it is
currently not at 0, and the real view currently
        # is at a level of 0
        BVSOLag = "\N"
        if BVExpCt <= eval(L['Minimum']): # Only perform count and
log count transaction if expected book view
        # inventory level is at or below the reorder point
            if stasku.realInvLev.amount == 0 &
stasku.realInvLev.amount != stasku.BVinv:
                BVSOevent = 1
                if eval(globalDict['LagRec']) == 1:
                    BVSOlag = now()-stasku.lastRVSO.pop()
                    stasku.monBVSO.observe(BVSOevent)
                if eval(globalDict['Plot']) == 1:
                    stasku.monBVSOFull.observe(BVSOevent)
            else: # conditions not met, do NOT record a book view
stockout
                BVSOevent = 0
                c = Count()
                c.runC(L, stasku)
                # Record transactional data for Count transactions
                stasku.transRec['Time'].append(now())
                stasku.transRec['Transaction Type'].append(4) #
"Transaction Type" = 4 indicates Count transaction
                stasku.transRec['Quantity'].append(0) # quantity has
no meaning for Count transaction
                stasku.transRec['BV Expected Count'].append(BVExpCt)

```

```

stasku.transRec['BV Actual Count'].append(BVActCt)
stasku.transRec['BV End Count'].append(BVEndCt)
stasku.transRec['RV Expected Count'].append(RVExpCt)
stasku.transRec['RV End Count'].append(RVEndCt)
stasku.transRec['Recorded'].append(1) # assume that
Count transaction is always recorded
stasku.transRec['CountEventTrue'].append(1)
stasku.transRec['RV StockoutEvent'].append(0)
stasku.transRec['BV StockoutEvent'].append(BVSOevent)
if eval(globalDict['LagRec']) == 1:
    stasku.transRec['BVSOLag'].append(BVSOLag)

class Count:
    def runC(self, L, stasku):
        # Increase or decrease the book view inventory level depending
        upon whether the real
        # inventor level is higher or lower than the book view inven-
        tory level, respectively.
        discrep = 0
        stasku.monNumCts.observe(1)
        if (stasku.realInvLev.amount < stasku.BVinv):
            stasku.absDiscrep.observe(stasku.BVinv-
stasku.realInvLev.amount) # Record absolute value of discrepancy
            discrep = stasku.realInvLev.amount-stasku.BVinv
            # If Count transaction creates BV stockout, record BVSO
start time
            if stasku.BVinv > 0 and discrep == -stasku.BVinv:
                stasku.tSinceBVSO = now()
            # Increase book view inventory level is real view is high-
er than the book view
            stasku.BVinv -= stasku.BVinv - stasku.realInvLev.amount
            stasku.monBV.observe(stasku.BVinv)
            if eval(globalDict['Plot']) == 1:
                stasku.monBVFull.observe(stasku.BVinv)
            elif (stasku.realInvLev.amount > stasku.BVinv):
                stasku.absDiscrep.observe(stasku.realInvLev.amount-
stasku.BVinv) # Record absolute value of discrepancy
                discrep = stasku.realInvLev.amount-stasku.BVinv
                # If Count transaction ends BV stockout, record BVSO dura-
tion time
                if stasku.BVinv <= 0 and discrep > -stasku.BVinv:
                    stasku.BVSOdurMon.observe(now() - stasku.tSinceBVSO)
                    stasku.tSinceBVSO = None
            # Decrease book view inventory level is real view is lower
than the book view
            stasku.BVinv += stasku.realInvLev.amount - stasku.BVinv

```

```

        stasku.monBV.observe(stasku.BVinv)
        if eval(globalDict['Plot']) == 1:
            stasku.monBVFull.observe(stasku.BVinv)
    stasku.p_estDemMon.observe(stasku.demSinceCt.total())
    stasku.p_estDiscrepMon.observe(discrep)
    stasku.demSinceCt.reset()
    stasku.CountRetMon.observe(stasku.retSinceCt.total())
    stasku.CountReplMon.observe(stasku.replSinceCt.total())
    stasku.retSinceCt.reset()
    stasku.replSinceCt.reset()

class Review(Process):
    def runRev(self, L, stasku, order):

        # Construct vector, 'IRt', of inter-review times for the cor-
        # responding day of the
        # week (e.g. IRt[0] is the time between the review on Monday,
        # and the next review).
        # Negative values in the vector indicate that no review occurs
        # on that day.
        R = ['R0', 'R1', 'R2', 'R3', 'R4', 'R5', 'R6']
        timeAdd = 1
        IRt = []
        for i in range(7):
            if eval(L[R[i]]) < 0:
                IRt.append(-1)
            else:
                if i < 6:
                    next = i+1
                else:
                    next = 0
                while eval(L[R[next]]) < 0:
                    timeAdd += 1
                    if next < 6:
                        next += 1
                    else:
                        next = 0
                IR = eval(L[R[next]])-eval(L[R[i]])+timeAdd
                IRt.append(IR)
                timeAdd = 1

        # Must start review cycle on a day at a scheduled review time
        startDay = d
        firstRevDay = d
        daysTilReview = 0
        startTime = eval(globalDict['Initial time bucket'])*tb

```

```

    # If review held on day that simulation started, and the re-
view time has not passed, start review cycle
    # in the review period that same day
    if eval(L[R[startDay]]) > 0 and startTime < ev-
al(L[R[startDay]]):
        yield hold, self, eval(L[R[startDay]])-startTime
    else:
        # If review time has passed for that day or there is no
review period scheduled for that day,
        # wait to start review cycle at least until midnight (t=0)
of the following day
        if startTime > eval(L[R[startDay]]):
            yield hold, self, 1-startTime
            if firstRevDay < 6:
                firstRevDay += 1
            else:
                firstRevDay = 0
        # If no review period occurs on the current day, cycle
through days of week until find the
        # day where the next review period occurs. If cycle
through the week and find no scheduled
        # review periods, print error message.
        while eval(L[R[firstRevDay]]) < 0:
            daysTilReview += 1
            if firstRevDay < 6:
                firstRevDay += 1
            else:
                firstRevDay = 0
        if firstRevDay == startDay:
            print "review does not occur on any day of the
week"
            exit()
        # After finding next review period, wait the calculated
number of days until that period,
        # and the fraction of the day past midnight on which the
review occurs
        yield hold, self, daysTilReview+eval(L[R[firstRevDay]])

    while True:
        # if want to count only products that are at or below re-
order point, schedule Count
        # event immediately before Review event (so that any inac-
curacies in the book view
        # level will be fixed before the order, so that the order
will be the correct size)
        if eval(L['CountLow']) == 1:

```



```

        scl = SourceCountLow()
        scl.runSCL(L, stasku)
        # Order may only be signaled if inventory level after de-
mand satisfied is less than or
        # equal to the reorder point
        if stasku.BVinv + sum(stasku.activeOrders) <= eval(L['Minimum']):
            reactivate(order)
            yield hold, self, IRt[d] # hold for time between reviews
(1 = daily)

class Order(Process):
    def runO(self,L,stasku):
        s = eval(L['Minimum']) # reorder point
        S = eval(L['Maximum']) # restock level
        while True:
            yield passivate, self # Wait for Review to reactivate
            orderSize = S - stasku.BVinv - sum(stasku.activeOrders)#
order up to restock level
            stasku.activeOrders.append(orderSize)
            r = Replenish()
            activate(r,r.runR(L,stasku,self))

class Replenish(Process):
    def runR(self,L,stasku,order):

        # Replenish Process Inputs
        lt = [eval(L['LT1']), eval(L['LT2']), eval(L['LT3']), eval(L['LT4']), eval(L['LT5']), eval(L['LT6']), eval(L['LT7'])]
        t = lt[d] # lead time (days), for v2, assume deterministic
lead time that is function of day of week
        k = eval(L['kill time']) # kill time (days)
        p = eval(L['Replenish Error probability']) # probability that
count is NOT called
        e = stasku.replRand.random()

        if t <= 0:
            print "lead time cannot be less than or equal to 0, re-
enter inputs"
            exit()
        else:
            # Replenish Process
            if t <= k: # lead time is less than or equal to kill time,
order will be filled

```

```

        yield hold, self, t # wait lead time
        BVExpCt = stasku.BVinv # book view expected count
        (prior to replenishment transaction)
        RVEExpCt = stasku.realInvLev.amount # real view ex-
        pected count (prior to replenishment transaction)
        RVEndCt = RVEExpCt + stasku.activeOrders[0] # real view
        end count (after replenishment transaction)
        BVSOlag = "\N"
        if e >= p: # Discrepancy between real and book view
        fixed before Replenishment occurs
            BVerror = RVEExpCt - BVExpCt
            BVActCt = BVExpCt + BVerror
            if BVActCt == 0 and BVExpCt > 0:
                BVSOevent = 1
                if eval(globalDict['LagRec']) == 1:
                    BVSOlag = now()-stasku.lastRVSO.pop()
            else:
                BVSOevent = 0
            c = Count()
            c.runC(L, stasku)
            CET = 1 # CountEventTrue
        else: # Discrepancy not fixed
            CET = 0 # CountEventTrue
            BVActCt = BVExpCt
            BVSOevent = 0

        stasku.monBVSO.observe(BVSOevent)
        if eval(globalDict['Plot']) == 1:
            stasku.monBVSOFull.observe(BVSOevent)

        BVEndCt = BVActCt + stasku.activeOrders[0] # book view
        end count (after replenishment transaction)
        # If replenishment transaction ends BV stockout, re-
        cord BVSO duration
        if BVActCt <= 0 and BVEndCt > 0:
            stasku.BVSOdurMon.observe(now()) -
stasku.tSinceBVSO)
            stasku.tSinceBVSO = None
        # If replenishment transaction ends RV stockout, re-
        cord RVSO duration
        if RVEExpCt == 0 and RVEndCt > 0:
            stasku.RVSOdurMon.observe(now()) -
stasku.tSinceRVSO)
            stasku.tSinceRVSO = None

        yield put, self, stasku.realInvLev,

```

```

stasku.activeOrders[0] # update real inventory level with received or-
der quantity
    if eval(globalDict['Plot']) == 1:

stasku.monRealFull.observe(stasku.realInvLev.amount)
    # for v2 assume that Replenishment is always recorded
accurately
    stasku.BVinv += stasku.activeOrders[0] # update book
view inventory level with received order quantity
    stasku.transQuant.observe(stasku.activeOrders[0]) #
record transaction quantity
    stasku.RVReplMon.observe(stasku.activeOrders[0])
    stasku.replSinceCt.observe(stasku.activeOrders[0])
    stasku.monBV.observe(stasku.BVinv)
    stasku.monNumRepl.observe(1)
    if eval(globalDict['Plot']) == 1:
        stasku.monBVFull.observe(stasku.BVinv)
        stasku.monNumReplFull.observe(1)
    # Record transactional data for Replenishment transac-
tions
    stasku.transRec['Time'].append(now())
    stasku.transRec['Transaction Type'].append(2) #
"Transaction Type" = 2 indicates Replenish (filled order) transaction

stasku.transRec['Quantity'].append(stasku.activeOrders[0]) # order
size of most recent order
    stasku.transRec['BV Expected Count'].append(BVExpCt)
    stasku.transRec['BV Actual Count'].append(BVActCt)
    stasku.transRec['BV End Count'].append(BVEndCt)
    stasku.transRec['RV Expected Count'].append(RVExpCt)
    stasku.transRec['RV End Count'].append(RVEndCt)
    stasku.transRec['Recorded'].append(1)
    stasku.transRec['CountEventTrue'].append(CET)
    stasku.transRec['RV StockoutEvent'].append(0)
    stasku.transRec['BV StockoutEvent'].append(BVSOevent)
    if eval(globalDict['LagRec']) == 1:
        stasku.transRec['BVSOLag'].append(BVSOlag)

stasku.activeOrders.pop(0) # remove order from active
orders list

    stasku.monKilled.observe(0)
    if eval(globalDict['Plot']) == 1:
        stasku.monKilledFull.observe(0)
else: # order is killed
    yield hold, self, k
    stasku.activeOrders.pop(0) # remove order from active

```

```

orders list
    stasku.monKilled.observe(1)
    if eval(globalDict['Plot']) == 1:
        stasku.monKilledFull.observe(1)

def main():
    global globalDict,d, t, tb, transLog, constVarsRec, statsRec
## Database Parameters-----
    # Define access parameters for the database here
    dbhost = "localhost"
    dbport = int("3306")
    dbuser = "*****"
    dbpass = "*****"

    #dbname = "david_validation"

## Filenames
    params_input_file = "v5input_validation_63940.csv"
    #params_input_file = "v5input_filename.csv"
    #global_input_file = "globalvarsinput_batch9.csv"
    global_input_file = "globalvarsinput_v5.csv"

    trans_output_file = "v5output_validation_63940.csv"
    params_output_file = "v5staskuvars_validation_63940.csv"
    stats_output_file = "v5stats_validation_63940.csv"

## Import of Simulation Parameters-----
    # Import & read file containing input per simulation
    f2 = open(global_input_file, "rU")
    dreader2 = csv.DictReader(f2)
    # Store global variables in dictionary, labeled according to first
row in .csv file

    globalDict = {}
    globalDict.update(dreader2.next())
    d = eval(globalDict['Initial day of week'])

    # Import & read file containing input per StaSKU
    f1 = open(params_input_file, "rU")
    dreader1 = csv.DictReader(f1)

    inputList = [] # initiate list of dictionaries of inputs

```

```

    for L in dreader1: # StaSKU-specific variables stored in dictionary,
        # labeled according to first row in .csv file
        alphaSum = eval(L['a0'])+eval(L['a1'])+eval(L['a2'])+eval(L['a3'])+\
            eval(L['a4'])+eval(L['a5'])+eval(L['a6'])+eval(L['a7'])+\
            eval(L['a8'])+eval(L['a9'])+eval(L['a10'])+eval(L['a11'])
        if math.fabs(alphaSum-1) > 0.01: # verify that daily demand
            fractions sum to 1
            print "Error: Demand Fraction do not sum to 1"
            exit()
        else:
            inputList.append(L) # Store dictionaries in list

f1.close()
f2.close()

of1 = open(trans_output_file, "wb")
transLog = csv.writer(of1, dialect='excel')
# Write column headers
transLog.writerow(('HashCode', 'Time', 'Transaction
Type', 'Quantity', \
    'BV Expected Count', 'BV Actual Count', 'BV End
Count', 'RV Expected Count', \
    'RV End Count', 'Recorded', 'CountEventTrue', 'RV
StockoutEvent', \
    'BV StockoutEvent', 'ID', 'BVSolag', 'Trial'))

# Output variables constant per Station SKU
of2 = open(params_output_file, "wb")
constVarsRec = csv.writer(of2, dialect='excel')
# Write column headers
constVarsRec.writerow(('ID', 'HashCode', 'FacilityStation', 'ItemID', 'Minimum', 'Maximum',
    'RunDateTime', 'Item
Class', 'InitialIOH', 'MeanDailyDemand', 'TB1frac', 'TB2frac', 'TB3frac', \
    'TB4frac', 'TB5frac', 'TB6frac', 'TB7frac', 'TB8frac', 'TB9frac', \
    'TB10frac', 'TB11frac', 'TB12frac', 'DemandError', 'InitialCount', 'inter-
count', \
    'InitialRe-
```

```

view', 'KillTime', 'ReplenishError', 'MaxSimTime', 'InitTimeBucket', \
      'NegBin-
Prob', 'InitialDayOfWeek', 'IRt1', 'IRt2', 'IRt3', 'IRt4', 'IRt5', \
      'IRt6', 'IRt7', 'LT1', 'LT2', 'LT3', 'LT4', 'LT5', 'LT6', 'LT7', 'ReturnProb', \
      'ReturnMean-
Time', 'Trial', 'Seed', 'R0', 'R1', 'R2', 'R3', 'R4', 'R5', 'R6', 'statTimeRange
', \
      'ReturnEr-
ror', 'CountLow', 'probCtOnLS', 'NumTrials'))

# Output simulation statistics
of3 = open(stats_output_file, "wb")
statsRec = csv.writer(of3, dialect='excel')
# Write column headers
stats-
Rec.writerow(('HashCode', 'StartPeriod', 'EndPeriod', 'RVDemand', 'LostSal
es', 'RVSales', \
      'BVSales', 'RVSO', 'BVSO', 'NumCts', 'NumRepl', 'RVPercAv', 'BVPercAv', \
      'AveRV-
Inv', 'AveBVInv', 'NumKilledOrd', 'RecProb', 'FillRate', 'RVSOrate', 'BVSOra
te', 'ID', \
      'TransQuant-
Sum', 'AbsDiscrepSum', 'CountDemSum', 'DiscrepSum', 'BVPercNeg', 'CountRetS
um', \
      'Coun-
tReplSum', 'RVReplSum', 'RVRetSum', 'BVRetSum', 'RVSOdur', 'BVSOdur', 'Trial
'))

## Model/Experiment -----
#   startTimeCounter = time.time()
#   # Station SKUs are simulated one at a time
lw = logWriter()
for L in inputList:
    hashCode = hash(hash(str(time.time()))+hash(str(L)))
    j = 1
    while j <= eval(globalDict['NumTrials']):
        initialize()
        dow = DayOfWeek() # Initialize day of week counter process
        activate(dow, dow.runDOW(L))
        TBproc = TimeBucket() # Initialize time bucket counter
process
        activate(TBproc, TBproc.runTB())
        #gaCo = GarbCollect() # Initialize periodic garbage col-
lection process

```

```

        #activate(gaCo,gaCo.runGC())

        seed = eval(globalDict['Seed'])+j
        trial = j
        # Pass input variables to new instance of StaSKU object
        s = StaSKU(L,seed,hashCode,trial)

        simulate(until=eval(globalDict['Run length']))

        ## Output of Results -----
        # Output transaction log
        lw.twrite(s)
        lw.swrite(s)

        print time.time(),"the time is now,", now(), "days"
        gc.collect()
        j += 1

    of1.close()
    of2.close()
    of3.close()

    # Export results to MySQL Database
    # Connect to MySQL server

db=MySQLdb.connect(host=dbhost,port=dbport,user=dbuser,passwd=dbpass,d
b=dbname)
    dbc=db.cursor();
    dbc.execute("""LOAD DATA LOW_PRIORITY LOCAL INFILE %s INTO TABLE
simruns_results
        FIELDS TERMINATED BY ',' LINES TERMINATED BY '\r\n' IGNORE 1
LINES;""", (trans_output_file,))
    dbc.execute("""LOAD DATA LOW_PRIORITY LOCAL INFILE %s INTO TABLE
simruns_params
        FIELDS TERMINATED BY ',' LINES TERMINATED BY '\r\n' IGNORE 1
LINES;""", (params_output_file,))
    dbc.execute("""LOAD DATA LOW_PRIORITY LOCAL INFILE %s INTO TABLE
simruns_stats
        FIELDS TERMINATED BY ',' LINES TERMINATED BY '\r\n' IGNORE 1
LINES
        (Hash-
Code,StartPeriod,EndPeriod,RVDemand,LostSales,RVSales,BVSO,BVS
O,NumCts,NumRepl,RVPerAv,BVPerAv,AveRVInv,AveBVInv,NumKilledOrd,RecP
rob,FillRate,RVSOrate,BVSOrate,ID,TransQuantSum,AbsDiscrepSum,CountDem
Sum,DiscrepSum,BVPerNeg,CountRetSum,CountReplSum,RVReplSum,RVRetSum,B
VRetSum,RVSOdur,BVSOdur,Trial);""", (stats_output_file,))

```

```

db.commit()
db.close()

class DayOfWeek(Process):
    def runDOW(self,L):
        global d
        # This process keeps track of which day of the week it is
        # d=0: Monday, d=1: Tuesday, d=2:Wednesday, d=3: Thursday,
d=4: Friday, d=5: Saturday, d=6: Sunday
        yield hold, self, 1-(eval(globalDict['Initial time buck-
et'])/12.0)
        if d < 6:
            d += 1
        elif d == 6:
            d = 0
        while True:
            yield hold, self, 1
            if d < 6:
                d += 1
            elif d == 6:
                d = 0

class TimeBucket(Process):
    def runTB(self):
        global tb, t
        # This process keeps track of what time bucket the simulation
is in
        tb = (1.0/12.0) # each "time bucket" is 2 hrs = 1/12 of a day
        t = eval(globalDict['Initial time bucket']) # initialize time
bucket
        while True:
            yield hold, self, tb
            # advance to next time bucket
            if t < 11:
                t += 1
            elif t == 11:
                t = 0

class GarbCollect(Process):
    def runGC(self):
        while True:
            # Call garbage collect periodically to reduce memory used,
but not so
            # often that it takes much longer to run
            yield hold, self, 364
            print time.time(),"the time is now,", now(), "days"

```



```
gc.collect()

if __name__ == '__main__': main()
```

Appendix E. Study of p for Two-Sided Inaccuracies

In order to characterize how p varies as the time since the last count t increases, it is useful to rewrite the cumulative distribution function $F_{N=Skellam(\lambda L+t\epsilon_2, t\epsilon_1)}$ using its characteristic function $\phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}$ through the inversion formula (Davies, 1973):

$$F_{N=Skellam(\lambda L+t\epsilon_2, t\epsilon_1)}(x+1) = P[N \leq x+1] = P[N < x] = \frac{1}{2} - \int_{-\pi}^{\pi} \operatorname{Re} \left[\frac{\phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}(u) e^{-iux}}{2\pi(1-e^{-iu})} \right] du$$

Where x is a non-negative integer.

The Skellam distribution has the following characteristic function $\phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}$, which is differentiable in t :

$$\begin{aligned} \phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}(u) &= \exp\left(-(\lambda L + t\epsilon_2 + t\epsilon_1) + (\lambda L + t\epsilon_2)e^{iu} + t\epsilon_1 e^{-iu}\right) \\ \frac{\partial \phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}}{\partial t}(u) &= \left[-(\epsilon_2 + \epsilon_1) + \epsilon_2 e^{iu} + \epsilon_1 e^{-iu}\right] \phi_{(\lambda L+t\epsilon_2, t\epsilon_1)}(u) \end{aligned}$$

Therefore, $F_{N=Skellam(\lambda L+t\epsilon_2, t\epsilon_1)}$ is also differentiable in t and therefore so is the function p .

$$\begin{aligned} \frac{\partial}{\partial t} F_{N=Skellam(\lambda L+t\varepsilon_2, t\varepsilon_1)}(x+1) &= -\frac{\partial}{\partial t} \int_{-\pi}^{\pi} \operatorname{Re} \left[\frac{\phi_{(\lambda L+t\varepsilon_2, t\varepsilon_1)}(u) e^{-iux}}{2\pi(1-e^{-iu})} \right] du \\ \frac{\partial}{\partial t} F_N(x+1) &= -\int_{-\pi}^{\pi} \operatorname{Re} \left[\frac{\frac{\partial \phi_{(\lambda L+t\varepsilon_2, t\varepsilon_1)}(u) e^{-iux}}{\partial t}}{2\pi(1-e^{-iu})} \right] du \\ \frac{\partial}{\partial t} F_N(x+1) &= -(\varepsilon_2 + \varepsilon_1) \left[F_N(x+1) - \frac{1}{2} \right] + \varepsilon_2 \left[F_N(x) - \frac{1}{2} \right] + \varepsilon_1 \left[F_N(x+2) - \frac{1}{2} \right] \\ \frac{\partial}{\partial t} F_N(x+1) &= \varepsilon_1 [F_N(x+2) - F_N(x+1)] - \varepsilon_2 [F_N(x+1) - F_N(x)] \\ \frac{\partial}{\partial t} F_N(x+1) &= \varepsilon_1 [f_N(x+2)] - \varepsilon_2 [f_N(x+1)] \end{aligned}$$

The sign of this derivative is a sufficient condition to conclude on the monotonicity of p :

$$\begin{aligned} \frac{\partial p}{\partial t}(t) &= \frac{1}{Q} \sum_{k=1}^Q \frac{\partial}{\partial t} F_{N=Skellam(\lambda L+t\varepsilon_2, t\varepsilon_1)}(R+k-1) \\ \frac{\partial p}{\partial t}(t) &= \frac{1}{Q} \sum_{k=1}^Q (\varepsilon_1 [f_N(R+k)] - \varepsilon_2 [f_N(R+k-1)]) \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial t} F_N(x+1) < 0 &\Leftrightarrow \varepsilon_1 [f_N(x+2)] < \varepsilon_2 [f_N(x+1)] \\ \Leftrightarrow \frac{f_N(x+2)}{f_N(x+1)} &< \frac{\varepsilon_2}{\varepsilon_1} \end{aligned}$$

Using the following change of variables, we can rewrite the inequality:

$$\delta = \frac{\varepsilon_1}{\varepsilon_2}, g = t\varepsilon_2, m = 2\sqrt{(\lambda L + g)g\delta}$$

$$\frac{e^{-(\lambda L + t\epsilon_2 + t\epsilon_1)} \left(\frac{\lambda L + t\epsilon_2}{t\epsilon_1} \right)^{\frac{x+2}{2}} I_{x+2} \left(2\sqrt{(\lambda L + t\epsilon_2)(t\epsilon_1)} \right)}{e^{-(\lambda L + t\epsilon_2 + t\epsilon_1)} \left(\frac{\lambda L + t\epsilon_2}{t\epsilon_1} \right)^{\frac{x+1}{2}} I_{x+1} \left(2\sqrt{(\lambda L + t\epsilon_2)(t\epsilon_1)} \right)} < \frac{\epsilon_2}{\epsilon_1}$$

$$\Leftrightarrow \frac{I_{x+2}(m)}{I_{x+1}(m)} < \frac{1}{\delta} \left(\frac{t\epsilon_1}{\lambda L + t\epsilon_2} \right)^{\frac{1}{2}} \Leftrightarrow \frac{I_{x+2}(m)}{I_{x+1}(m)} < \frac{2g}{m}$$

We can express g as a function of the other variables:

$$m = 2\sqrt{(\lambda L + g)g\delta}$$

$$g^2 + \lambda Lg - \frac{m^2}{4\delta} = 0$$

$$g = \frac{-\lambda L + \sqrt{(\lambda L)^2 + \frac{m^2}{\delta}}}{2} = \frac{\lambda L}{2} \left(\sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}} - 1 \right)$$

$$g = \frac{\lambda L}{2} \frac{m^2}{(\lambda L)^2 \delta} \frac{1}{1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}}} = \frac{m^2}{2(\lambda L)\delta} \frac{1}{1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}}}$$

The ratio of modified Bessel functions of the first kind is bound by a lower and upper bound: (Amos, 1974):

$$l(x, m) = \frac{m}{x + 2 + \sqrt{m^2 + (x + 2)^2}} \leq \frac{I_{x+2}(m)}{I_{x+1}(m)} \leq u(x, m) = \frac{m}{x + 1 + \sqrt{m^2 + (x + 3)^2}}$$

The upper bound yields a sufficient condition for p decreasing:

$$u(x,m) < \frac{m}{(\lambda L)\delta} \frac{1}{1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}}} = \frac{2g}{m}, \forall x \in [R-1, R+Q-1], \forall m > 0 \Rightarrow \frac{\partial p}{\partial t}(t) < 0, \forall t > 0$$

Since $u(x,m)$ is decreasing in x :

$$u(R-1,m) < \frac{m}{(\lambda L)\delta} \frac{1}{1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}}}, \forall m > 0 \Rightarrow \frac{\partial p}{\partial t}(t) < 0, \forall t > 0$$

$$\Leftrightarrow u(R-1,m) < \frac{m}{(\lambda L)\delta} \frac{1}{1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}}}, \forall m > 0$$

$$\Leftrightarrow q(m) = R + \sqrt{m^2 + (R+2)^2} - (\lambda L)\delta \left[1 + \sqrt{1 + \frac{m^2}{(\lambda L)^2 \delta}} \right] > 0, \forall m > 0$$

$$\Leftrightarrow q(m) = R - (\lambda L)\delta + m \left[\sqrt{1 + \frac{(R+2)^2}{m^2}} - \sqrt{\delta + \frac{(\lambda L)^2 \delta^2}{m^2}} \right] > 0, \forall m > 0$$

Since $\delta \leq 1$, we have the desired result:

$$R > \lambda L \delta \Rightarrow q(m) > 0, \forall m > 0 \Rightarrow \frac{\partial p}{\partial t}(t) < 0, \forall t > 0$$

Appendix F. Study of the Convexity of MAPL

To examine the convexity of the expected in-stock probability, we consider the lower bound on the expected in-stock probability obtained by Morey (1985). He defines the Minimum Actual Protection Level (MAPL) as the probability of the sum of the demand D_L and *maximum* inventory inaccuracy during the period M_t being below the reorder point B , as a function of the number of periods since the last count t :

$$MAPL(B, t) = \Pr[D_L + M_t \leq E[D_L] + B]$$

Under the assumptions of approximately normal lead time demand and independence of lead time demand and maximum inventory inaccuracy between counts⁶⁶, he obtains an analytical expression for the Minimum Actual Protection Level (MAPL):

$$s_{MAPL}(t) = MAPL(B, t) = 2 \cdot \Phi\left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\epsilon^2}}\right) - \Phi\left(\frac{B}{\sigma_{DL}}\right)$$

Because this expression is a lower bound (hence the use of the word “Minimum”) on the average expected in-stock probability, we have:

⁶⁶ Morey treats the case of non-zero mean inventory inaccuracies by including their mean in the demand, which makes the problem tractable as long as the maximum inventory inaccuracies during the period can still be considered independent of lead-time demand.

$$s_{MAPL}(t) \leq s(t) \Rightarrow S_{MAPL}(T) = \frac{1}{\mu} E[T s_{MAPL}(T)] \leq S(T)$$

We now consider s_{MAPL} as an approximation of s . We consider $r_{MAPL}: t \rightarrow t s_{MAPL}(t)$ and verify that it is strictly concave over $[0, +\infty)$, by observing that r_{MAPL} is twice differentiable over this interval and calculating its second derivative:

$$\begin{aligned} r_{MAPL}(t) &= t \left[2 \cdot \Phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) - \Phi \left(\frac{B}{\sigma_{DL}} \right) \right] \\ r_{MAPL}'(t) &= \left[2 \cdot \Phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) - \Phi \left(\frac{B}{\sigma_{DL}} \right) \right] - t \cdot B\sigma_\varepsilon^2 (\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{-3/2} \phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) \\ r_{MAPL}''(t) &= -B\sigma_\varepsilon^2 (\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{-3/2} \phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) - B\sigma_\varepsilon^2 (\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{-3/2} \phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) \\ &\quad - t \cdot \left[\frac{1}{2} B^3 \sigma_\varepsilon^4 (\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{-7/2} \phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) \right] \\ r_{MAPL}''(t) &= -B\sigma_\varepsilon^2 (\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{-7/2} \phi \left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}} \right) \left[2(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^2 + tB^2\sigma_\varepsilon^2 \right] < 0 \quad \forall t \geq 0 \end{aligned}$$

Therefore, the results of section 1 apply, i.e. the optimal service level is achieved for equally spaced counts:

$$ISP_{MAPL}(T) = \frac{1}{\mu} E[T s_{MAPL}(T)] \leq s_{MAPL}(\mu)$$

We examine the convexity of s_{MAPL} :

$$\begin{aligned}
s_{MAPL}'(t) &= -\frac{B\sigma_\varepsilon^2}{(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{3/2}} \cdot \varphi\left(\frac{B}{\sqrt{\sigma_{DL}^2 + t\sigma_\varepsilon^2}}\right) = -\frac{B\sigma_\varepsilon^2}{\sqrt{2\pi}(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{3/2}} \cdot \exp\left(-\frac{1}{2} \frac{B^2}{\sigma_{DL}^2 + t\sigma_\varepsilon^2}\right) \\
s_{MAPL}''(t) &= \frac{1}{\sqrt{2\pi}} \frac{3}{2} \sigma_\varepsilon^2 \frac{B\sigma_\varepsilon^2}{(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{5/2}} \cdot \exp\left(-\frac{1}{2} \frac{B^2}{\sigma_{DL}^2 + t\sigma_\varepsilon^2}\right) \\
&\quad - \frac{B\sigma_\varepsilon^2}{\sqrt{2\pi}(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{3/2}} \cdot \left(\frac{1}{2} \frac{B^2\sigma_\varepsilon^2}{(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^2}\right) \exp\left(-\frac{1}{2} \frac{B^2}{\sigma_{DL}^2 + t\sigma_\varepsilon^2}\right) \\
s_{MAPL}'''(t) &= \frac{1}{2\sqrt{2\pi}} \frac{B\sigma_\varepsilon^4}{2(\sigma_{DL}^2 + t\sigma_\varepsilon^2)^{7/2}} \exp\left(-\frac{1}{2} \frac{B^2}{\sigma_{DL}^2 + t\sigma_\varepsilon^2}\right) \left[3(\sigma_{DL}^2 + t\sigma_\varepsilon^2) - B^2\right]
\end{aligned}$$

Therefore, s_{MAPL} is convex iff:

$$\begin{aligned}
s_{MAPL}'''(t) &\geq 0 \\
3(\sigma_{DL}^2 + t\sigma_\varepsilon^2) - B^2 &\geq 0 \\
t &\geq \frac{\frac{1}{3}B^2 - \sigma_{DL}^2}{\sigma_\varepsilon^2} = t_{critical}(B, \sigma_{DL}^2, \sigma_\varepsilon^2)
\end{aligned}$$

The strict convexity of s_{MAPL} for all non-negative values of t is equivalent to:

$$\begin{aligned}
t_{critical}(B, \sigma_{DL}^2, \sigma_\varepsilon^2) &< 0 \\
B &< \sqrt{3} \cdot \sigma_{DL}
\end{aligned}$$

Under the assumption of normally distributed lead time demand, $t_{critical}$ can be expressed as a function of the target Type I service level α , i.e. the probability of not stocking out during a replenishment cycle in the absence of inventory inaccuracy:

$$B = k\sigma_{DL} = \Phi^{-1}(\alpha)\sigma_{DL}$$

$$t_{critical}(\alpha, \sigma_{DL}^2, \sigma_{\varepsilon}^2) = \frac{\frac{1}{3}B^2 - \sigma_{DL}^2}{\sigma_{\varepsilon}^2} = \frac{\sigma_{DL}^2}{\sigma_{\varepsilon}^2} \left(\frac{(\Phi^{-1}(\alpha))^2 - 3}{3} \right)$$

This analysis shows that the convexity of Morey's MAPL is guaranteed for all distributions if the target service level is below , as well as all distributions of T which take values above $t_{critical}$. In the next section, we consider different numerical applications to show that in most operational contexts the Minimum Actual Protection Level (MAPL) exhibits diminishing marginal decreases and that therefore $s_{MAPL}(\mu_{eq})$ is a lower bound on the service level.

Appendix G. Sensitivity Analysis for Parameter K

α	CV	K	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.80						
	0.10					
		0.10	0.7775	0.0000	0.0002	0.0223
		0.25	0.7480	0.0000	0.0004	0.0516
		0.50	0.7073	0.0001	0.0006	0.0920
		1.00	0.6472	0.0003	0.0007	0.1518
	0.25					
		0.10	0.7764	0.0001	0.0012	0.0223
		0.25	0.7456	0.0004	0.0024	0.0516
		0.50	0.7036	0.0010	0.0034	0.0920
		1.00	0.6421	0.0021	0.0040	0.1518
	0.50					
		0.10	0.7725	0.0005	0.0047	0.0223
		0.25	0.7374	0.0018	0.0091	0.0516
		0.50	0.6909	0.0043	0.0129	0.0920
		1.00	0.6253	0.0081	0.0149	0.1518
	1.00					
		0.10	0.7577	0.0019	0.0181	0.0223
		0.25	0.7080	0.0070	0.0334	0.0516
		0.50	0.6482	0.0147	0.0451	0.0920
		1.00	0.5730	0.0241	0.0511	0.1518
	1.50					
		0.10	0.7353	0.0051	0.0373	0.0223
		0.25	0.6681	0.0155	0.0648	0.0516
		0.50	0.5966	0.0276	0.0838	0.0920
		1.00	0.5169	0.0388	0.0925	0.1518

Table 9.1: Service Level Breakdown for different values of K, ($\alpha = 0.80$)

α	CV	K	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.90						
0.1						
		0.10	0.8780	0.0000	0.0002	0.0217
		0.25	0.8478	0.0000	0.0004	0.0517
		0.50	0.8038	0.0001	0.0007	0.0954
		1.00	0.7340	0.0003	0.0009	0.1648
0.25						
		0.10	0.8769	0.0001	0.0012	0.0217
		0.25	0.8454	0.0003	0.0026	0.0517
		0.50	0.7996	0.0009	0.0041	0.0954
		1.00	0.7278	0.0022	0.0052	0.1648
0.5						
		0.10	0.8731	0.0004	0.0048	0.0217
		0.25	0.8367	0.0015	0.0101	0.0517
		0.50	0.7853	0.0040	0.0154	0.0954
		1.00	0.7071	0.0085	0.0195	0.1648
1						
		0.10	0.8580	0.0014	0.0189	0.0217
		0.25	0.8046	0.0060	0.0377	0.0517
		0.50	0.7352	0.0146	0.0549	0.0954
		1.00	0.6406	0.0273	0.0672	0.1648
1.5						
		0.10	0.8344	0.0041	0.0397	0.0217
		0.25	0.7589	0.0148	0.0746	0.0517
		0.50	0.6711	0.0300	0.1035	0.0954
		1.00	0.5658	0.0469	0.1224	0.1648

Table 9.2: Service Level Breakdown for different values of K, ($\alpha = 0.90$)

α	CV	K	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.95						
0.1						
		0.10	0.9330	0.0000	0.0002	0.0168
		0.25	0.9084	0.0000	0.0004	0.0412
		0.50	0.8700	0.0001	0.0007	0.0793
		1.00	0.8040	0.0002	0.0010	0.1448
0.25						
		0.10	0.9322	0.0000	0.0010	0.0168
		0.25	0.9063	0.0002	0.0023	0.0412
		0.50	0.8662	0.0006	0.0039	0.0793
		1.00	0.7979	0.0017	0.0056	0.1448
0.5						
		0.10	0.9290	0.0002	0.0039	0.0168
		0.25	0.8989	0.0009	0.0089	0.0412
		0.50	0.8531	0.0027	0.0150	0.0793
		1.00	0.7772	0.0070	0.0210	0.1448
1						
		0.10	0.9168	0.0006	0.0158	0.0168
		0.25	0.8707	0.0036	0.0344	0.0412
		0.50	0.8052	0.0110	0.0545	0.0793
		1.00	0.7077	0.0246	0.0729	0.1448
1.5						
		0.10	0.8970	0.0022	0.0340	0.0168
		0.25	0.8282	0.0106	0.0700	0.0412
		0.50	0.7400	0.0257	0.1050	0.0793
		1.00	0.6251	0.0461	0.1341	0.1448

Table 9.3: Service Level Breakdown for different values of K, ($\alpha = 0.95$)

α	CV	K	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.98						
0.1						
		0.10	0.9697	0.0000	0.0001	0.0102
		0.25	0.9535	0.0000	0.0003	0.0262
		0.50	0.9259	0.0000	0.0005	0.0536
		1.00	0.8726	0.0001	0.0009	0.1064
0.25						
		0.10	0.9691	0.0000	0.0006	0.0102
		0.25	0.9521	0.0000	0.0017	0.0262
		0.50	0.9230	0.0002	0.0032	0.0536
		1.00	0.8673	0.0009	0.0054	0.1064
0.5						
		0.10	0.9672	0.0001	0.0026	0.0102
		0.25	0.9470	0.0002	0.0066	0.0262
		0.50	0.9128	0.0010	0.0125	0.0536
		1.00	0.8491	0.0043	0.0202	0.1064
1						
		0.10	0.9592	-0.0001	0.0107	0.0102
		0.25	0.9264	0.0008	0.0265	0.0262
		0.50	0.8736	0.0055	0.0474	0.0536
		1.00	0.7843	0.0178	0.0715	0.1064
1.5						
		0.10	0.9456	0.0001	0.0240	0.0102
		0.25	0.8929	0.0045	0.0564	0.0262
		0.50	0.8151	0.0170	0.0944	0.0536
		1.00	0.7009	0.0389	0.1338	0.1064

Table 9.4: Service Level Breakdown for different values of K, ($\alpha = 0.98$)

Note: Small negative differences between the simulated Actual Service Level and the Lower Bound may be due to non-convexity of the service level function for these parameters, as described in Chapter 7.2.1.

α	CV	K	SL Lower Bound	Physical SL - SL Lower Bound	SL Loss due to Count Variability	SL Loss due to Inventory Inaccuracy
0.99						
0.1						
		0.10	0.9834	0.0000	0.0001	0.0065
		0.25	0.9724	0.0000	0.0002	0.0175
		0.50	0.9521	0.0000	0.0004	0.0375
		1.00	0.9092	0.0000	0.0008	0.0800
0.25						
		0.10	0.9830	0.0000	0.0004	0.0065
		0.25	0.9713	0.0000	0.0012	0.0175
		0.50	0.9499	0.0000	0.0026	0.0375
		1.00	0.9047	0.0004	0.0049	0.0800
0.5						
		0.10	0.9817	0.0000	0.0018	0.0065
		0.25	0.9677	-0.0001	0.0049	0.0175
		0.50	0.9420	0.0002	0.0103	0.0375
		1.00	0.8891	0.0024	0.0186	0.0800
1						
		0.10	0.9763	-0.0004	0.0075	0.0065
		0.25	0.9525	-0.0005	0.0206	0.0175
		0.50	0.9100	0.0021	0.0404	0.0375
		1.00	0.8308	0.0124	0.0669	0.0800
1.5						
		0.10	0.9667	-0.0007	0.0174	0.0065
		0.25	0.9260	0.0010	0.0455	0.0175
		0.50	0.8590	0.0106	0.0829	0.0375
		1.00	0.7509	0.0318	0.1273	0.0800

Table 9.5: Service Level Breakdown for different values of K, ($\alpha = 0.99$)

Note: Small negative differences between the simulated Actual Service Level and the Lower Bound may be due to non-convexity of the service level function for these parameters, as described in Chapter 7.2.1.

Appendix H. Count Variability in Hospital Stations

Row Labels	Average of Mean Inter-Count Time	Average of C.V.	Number of Station-SKUs	P-Value < 0.01	P-Value < 0.05	p-value < 0.10
01.S_5W	12.9	1.50	197	42	71	84
01.S_6C	10.2	1.00	197	78	106	123
01.S_6E	29.3	1.59	37	1	3	8
01.S_6SE	24.8	1.42	33	0	2	3
01.S_6W	15.6	1.06	170	24	37	61
01.S_7C	20.1	0.96	104	1	8	14
01.S_7E	28.8	1.38	38	0	2	5
01.S_7SE	26.1	1.49	32	0	0	3
01.S_7W	22.7	0.96	136	5	14	19
01.S_AM_ANES2	3.5	0.96	35	3	5	7
01.S_AMB_SURG	20.0	1.06	76	12	35	42
01.S_ANES	11.6	1.58	122	42	51	60
01.S_ANGIO_1	32.4	1.41	12	4	6	7
01.S_ANGIO_2	37.0	1.01	4	1	1	2
01.S_CCU_1	24.6	1.82	48	9	10	12
01.S_CCU_2	16.7	1.53	123	7	15	24
01.S_CL_A	7.6	0.78	476	14	29	47
01.S_CL_B	8.2	0.69	583	8	22	34
01.S_CL_C	11.0	0.90	409	26	42	55
01.S_CL_SR_1	24.9	1.40	412	31	73	103
01.S_CL_SR_2	35.5	1.32	21	2	6	7
01.S_ED_IV	10.0	2.24	21	3	4	7
01.S_ED_MINOR	21.1	1.36	73	10	20	27
01.S_ED_NS_1	5.4	1.56	4	0	0	0
01.S_ED_NS_2	24.0	1.03	2	1	2	2
01.S_ED_RM_5	45.0	1.50	1	0	1	1
01.S_ED_TELE	19.9	2.00	21	7	10	10
01.S_ED_TRA_1	31.4	1.56	11	2	3	4
01.S_ED_TRA_2	33.4	1.35	5	0	0	0
01.S_ED_UTIL	17.9	1.42	114	11	19	29
01.S_ENDO_1	34.9	1.21	9	1	2	3
01.S_ENDO_2	34.7	1.21	11	1	2	2
01.S_ENDO_3	36.0	1.19	8	1	2	2

Row Labels	Average of Mean Inter-Count Time	Average of C.V.	Number of Station-SKUs	P-Value < 0.01	P-Value < 0.05	p-value < 0.10
01.S_ENDO_4	32.2	1.30	13	1	1	4
01.S_ENDO_5	15.1	1.82	51	21	24	31
01.S_ENDO_6	28.0	1.53	19	1	2	3
01.S_ENDO_7	30.7	1.74	11	3	6	9
01.S_ENDO_RR	16.6	1.51	47	5	13	15
01.S_INR	23.4	0.97	1	0	1	1
01.S_IV_THERA	17.3	1.83	20	0	2	2
01.S_LITHO	26.8	2.41	1	0	0	0
01.S_MICU_1	24.1	1.79	38	4	7	11
01.S_MICU_2	18.6	1.82	112	8	19	29
01.S_N_ANES	28.4	1.29	17	1	6	6
01.S_N_ED	24.1	0.92	42	2	5	8
01.S_N_ED_ORT	24.5	0.90	44	3	8	9
01.S_N_ENDO_1	30.8	1.28	1	0	0	0
01.S_N_ENDO_2	39.3	1.05	1	0	0	1
01.S_N_ENDO_3	21.8	1.12	6	1	1	3
01.S_N_INPT	27.6	0.92	31	2	8	14
01.S_N_OR_1	40.9	1.42	1	0	0	0
01.S_N_OR_BLD	32.9	1.26	3	1	1	1
01.S_N_OR_CR2	31.4	0.94	1	0	0	0
01.S_N_OR_EYE	21.2	1.19	1	1	1	1
01.S_N_OR_ORT	36.8	1.05	5	0	2	4
01.S_N_OR_PK1	24.7	1.39	36	2	5	8
01.S_N_OR_PK2	23.1	1.33	14	0	1	2
01.S_N_PACU	21.1	1.51	37	3	11	13
01.S_N_TX_RM	11.6	0.83	1	0	0	0
01.S_OR_CAR_1	14.9	0.94	316	96	142	167
01.S_OR_CR_1	26.5	0.94	6	3	4	5
01.S_OR_CR_2	25.6	1.16	42	21	24	30
01.S_OR_ENT	10.1	0.83	6	1	3	5
01.S_OR_GEN	24.1	0.92	92	48	55	62
01.S_OR_GYN	24.3	0.98	93	62	79	82
01.S_OR_LAP	20.2	1.02	66	48	57	57
01.S_OR_LAP_2	23.9	1.04	104	61	73	77
01.S_OR_ORTD	22.6	1.45	51	8	13	19
01.S_OR_PLS	27.2	1.00	3	1	2	2
01.S_OR_SR	17.8	1.00	39	24	27	27
01.S_OR_STK_2	29.8	1.00	27	2	6	11
01.S_OR_SUT	19.3	0.94	129	100	114	119
01.S_OR_URO	12.4	0.84	121	45	69	76
01.S_OR_URO_2	11.3	0.92	52	24	29	35
01.S_PAC_SYR2	15.5	1.15	13	2	3	4

Row Labels	Average of Mean Inter-Count Time	Average of C.V.	Number of Station-SKUs	P-Value < 0.01	P-Value < 0.05	p-value < 0.10
01.S_PACU	13.3	1.13	105	23	42	52
01.S_PACU_2	13.2	0.98	39	7	16	20
01.S_PACU_3	11.0	1.16	27	14	18	18
01.S_PACU_SYR	14.9	1.35	20	4	5	7
01.S_SICU_A_1	18.9	1.01	102	10	22	25
01.S_SICU_A_2	19.7	1.09	90	5	17	28
01.S_SICU_A_3	14.3	0.93	6	0	0	0
01.S_SICU_A_4	12.1	1.04	6	0	0	0
01.S_SICU_B_1	19.3	1.01	99	6	19	26
01.S_SICU_B_2	22.0	1.12	90	12	23	28
01.S_SICU_B_3	12.9	1.08	6	0	0	1
01.S_SICU_B_4	14.4	0.98	6	0	0	0
Grand Total	16.4	1.10	5855	1033	1589	1965

Table 9.6: Average of Mean Inter-Count Time, Average of C.V., Number of SKUs, and Number of SKUs for which we reject the null at the 0.01, 0.05 and 0.10 significance levels

Note: This table presents Station-SKUs for which there was sufficient data to conduct the Fischer test of independence. This required more than ten counts during the study period and sufficient transaction data to build a histogram of the steady state distribution of the book inventory and perform the Fischer test of independence. 255 Station-SKUs (6,110 – 5,855) were omitted because of insufficient transaction data.

Appendix I. Sample Panel Simulation Results

ConfigName	ItemID	Accuracy	InterCo	Dyami	Recordin										CVsaw	CVsdt	Replave	Replst	RVmva	RVmvt	BVMva	BVMvt				
					g	Minimu	Repr	Flavr	Rstsd	RSV0A0	RSV0S0	BVS0A0	BVS0S0	RVPAa0									RVPASt	BVPAa0	BVPASt	
41	64040	Acc	RedVarIC	COR	COR100	0.94792	16	0	1.000	0.000	0.001	0.003	0.001	0.003	0.000	0.000	1.000	0.001	0.000	0.000	0.460	0.019	29.634	0.412	30.509	0.378
42	64040	Acc	RedVarIC	COR	COR50	0.94792	16	0.5	1.000	0.001	0.004	0.008	0.002	0.005	1.000	0.001	1.000	0.000	0.364	0.038	0.458	0.019	29.029	0.618	30.744	0.382
43	64040	Acc	GRIC	DIP	COR	0.97396	Adjust16	1	1.000	0.001	0.004	0.011	0.000	0.000	1.000	0.001	1.000	0.000	0.145	0.053	0.460	0.017	30.674	1.314	33.282	1.440
44	64040	Acc	GRIC	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.605	0.058	0.461	0.018	31.615	0.528	32.049	0.533
45	64040	Acc	GRIC	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.602	0.050	0.460	0.018	31.615	0.528	32.049	0.533
46	64040	Acc	GRIC	noDIP	COR	0.97396	16	1	0.996	0.023	0.008	0.015	0.001	0.004	0.996	0.025	1.000	0.001	0.145	0.053	0.455	0.023	28.454	1.511	31.057	0.846
47	64040	Acc	GRIC	noDIP	COR100	0.97396	16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.604	0.057	0.459	0.017	30.352	0.441	30.795	0.374
48	64040	Acc	GRIC	noDIP	COR50	0.97396	16	0.5	1.000	0.000	0.001	0.005	0.000	0.002	1.000	0.000	1.000	0.000	0.374	0.046	0.459	0.018	29.936	0.527	30.846	0.376
49	64040	Acc	HiHMean	DIP	COR	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.582	0.059	0.460	0.018	31.718	0.429	31.928	1.488
50	64040	Acc	HiHMean	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.593	0.056	0.461	0.019	31.635	0.546	32.073	0.571
51	64040	Acc	HiHMean	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.362	0.044	0.462	0.018	31.476	0.580	32.410	0.626
52	64040	Acc	HiHMean	noDIP	COR	0.97396	16	1	0.994	0.030	0.014	0.022	0.001	0.003	0.994	0.031	1.000	0.000	0.132	0.050	0.456	0.024	28.208	1.546	31.012	0.462
53	64040	Acc	HiHMean	noDIP	COR100	0.97396	16	0	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.592	0.055	0.460	0.017	30.368	0.600	30.808	0.367
54	64040	Acc	HiHMean	noDIP	COR50	0.97396	16	0.5	1.000	0.000	0.001	0.003	0.000	0.000	1.000	0.000	1.000	0.000	0.362	0.045	0.460	0.017	29.896	0.519	30.860	0.372
55	64040	Acc	HiHMean	DIP	COR	0.97396	Adjust16	1	1.000	0.000	0.000	0.002	0.000	0.000	1.000	0.000	1.000	0.000	0.125	0.012	0.461	0.018	30.891	0.639	32.580	0.535
56	64040	Acc	HiHMean	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.586	0.022	0.461	0.018	31.522	0.581	31.956	0.686
57	64040	Acc	HiHMean	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.002	0.000	0.000	1.000	0.000	1.000	0.000	0.121	0.000	0.459	0.017	29.706	0.810	32.564	0.560
58	64040	Acc	HiHMean	noDIP	COR	0.97396	16	1	1.000	0.001	0.002	0.006	0.000	0.000	1.000	0.000	1.000	0.000	0.125	0.012	0.459	0.018	29.363	0.660	31.052	0.388
59	64040	Acc	HiHMean	noDIP	COR100	0.97396	16	0	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.585	0.023	0.460	0.018	30.350	0.466	30.797	0.377
60	64040	Acc	HiHMean	noDIP	COR50	0.97396	16	0.5	1.000	0.000	0.001	0.004	0.000	0.002	1.000	0.000	1.000	0.000	0.355	0.038	0.459	0.017	29.966	0.598	30.889	0.412
61	64040	Acc	Histic	DIP	COR	0.97396	Adjust16	1	0.999	0.033	0.010	0.016	0.001	0.004	0.999	0.022	1.000	0.000	0.121	0.000	0.459	0.017	29.706	0.810	32.564	0.560
62	64040	Acc	Histic	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.582	0.018	0.461	0.018	31.727	0.701	32.163	0.694
63	64040	Acc	Histic	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.001	0.003	0.000	0.000	1.000	0.000	1.000	0.000	0.351	0.043	0.461	0.016	31.381	0.777	32.405	0.750
64	64040	Acc	Histic	noDIP	COR	0.97396	16	1	0.998	0.003	0.014	0.017	0.001	0.004	0.999	0.022	1.000	0.000	0.121	0.000	0.457	0.019	29.257	0.996	31.116	0.399
65	64040	Acc	Histic	noDIP	COR100	0.97396	16	0	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.583	0.019	0.460	0.018	30.321	0.627	31.381	0.388
66	64040	Acc	Histic	noDIP	COR50	0.97396	16	0.5	1.000	0.001	0.002	0.006	0.001	0.003	1.000	0.001	1.000	0.000	0.351	0.042	0.460	0.018	29.781	0.638	30.827	0.389
67	64040	Acc	LowMean	DIP	COR	0.97396	Adjust16	1	1.000	0.001	0.002	0.009	0.000	0.000	1.000	0.001	1.000	0.000	0.156	0.057	0.460	0.018	30.754	1.258	33.175	1.334
68	64040	Acc	LowMean	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.616	0.062	0.461	0.018	31.611	0.537	32.025	0.544
69	64040	Acc	LowMean	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.385	0.050	0.461	0.019	31.429	0.588	32.314	0.659
70	64040	Acc	LowMean	noDIP	COR	0.97396	16	0	0.988	0.166	0.029	0.047	0.001	0.004	0.997	0.164	1.000	0.001	0.149	0.086	0.460	0.023	31.633	0.021	31.433	0.427
71	64040	Acc	LowMean	noDIP	COR100	0.97396	16	0	1.000	0.000	0.000	0.002	0.000	0.002	1.000	0.000	1.000	0.000	0.617	0.062	0.461	0.018	30.382	0.438	30.812	0.375
72	64040	Acc	LowMean	noDIP	COR50	0.97396	16	0.5	1.000	0.000	0.001	0.005	0.000	0.002	1.000	0.000	1.000	0.000	0.385	0.049	0.460	0.018	29.950	0.555	30.858	0.381
73	64040	Acc	LowMean	DIP	COR	0.97396	Adjust16	1	1.000	0.000	0.000	0.002	0.000	0.000	1.000	0.000	1.000	0.000	0.147	0.014	0.460	0.017	30.898	0.525	32.398	0.589
74	64040	Acc	LowMean	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.376	0.037	0.459	0.017	30.867	0.600	30.907	0.441
75	64040	Acc	LowMean	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.376	0.038	0.460	0.018	31.287	0.573	32.104	0.529
76	64040	Acc	LowMean	noDIP	COR	0.97396	16	1	1.000	0.000	0.001	0.004	0.000	0.000	1.000	0.000	1.000	0.000	0.147	0.014	0.459	0.018	29.513	0.675	31.013	0.437
77	64040	Acc	LowMean	noDIP	COR100	0.97396	16	0	1.000	0.000	0.000	0.002	0.000	0.002	1.000	0.000	1.000	0.000	0.607	0.023	0.460	0.018	30.336	0.458	30.765	0.342
78	64040	Acc	LowMean	noDIP	COR50	0.97396	16	0.5	1.000	0.000	0.001	0.005	0.000	0.002	1.000	0.000	1.000	0.000	0.376	0.037	0.459	0.017	30.867	0.600	30.907	0.441
79	64040	Acc	RedVarIC	DIP	COR	0.97396	Adjust16	1	1.000	0.000	0.000	0.002	0.000	0.000	1.000	0.000	1.000	0.000	0.135	0.013	0.460	0.018	30.861	0.518	32.448	0.517
80	64040	Acc	RedVarIC	DIP	COR100	0.97396	Adjust16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.596	0.022	0.461	0.018	31.555	0.505	31.983	0.537
81	64040	Acc	RedVarIC	DIP	COR50	0.97396	Adjust16	1	1.000	0.000	0.000	0.002	0.000	0.000	1.000	0.000	1.000	0.000	0.365	0.040	0.460	0.019	31.308	0.466	32.178	0.521
82	64040	Acc	RedVarIC	noDIP	COR	0.97396	16	1	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	1.000	0.000	0.135	0.013	0.460	0.018	29.627	0.574	30.921	

ConfigIndex	ItemID	Accuracy	InterCo	Dynam	Recordin													Ctsave	Ctsstd	Replave	Replstd	RVInva	RVInvst	BVInva	BVInvst	
					g	Minimu	ReplErr	FRave	FRstd	RVS0av	RVS0st	BVS0av	BVS0st	RVPAav	RVPAst	BVPAav	BVPAst									
59	69111	IAcc	HighMean	noDIP	COR100	0.84805	20	0	0.804	0.050	0.296	0.070	0.295	0.070	0.930	0.026	0.966	0.009	1.852	0.099	1.194	0.095	21.628	0.828	23.276	0.361
60	69111	IAcc	HighMean	noDIP	COR50	0.84805	20	0.5	0.799	0.050	0.302	0.072	0.260	0.064	0.929	0.025	0.967	0.009	1.251	0.075	1.183	0.097	21.386	0.831	23.350	0.348
61	69111	IAcc	HistC	DIP	COR0	0.84805	AdjIust2C	1	0.818	0.048	0.281	0.064	0.171	0.048	0.937	0.022	0.973	0.008	1.121	0.000	1.255	0.102	23.293	1.077	26.348	1.053
62	69111	IAcc	HistC	DIP	COR100	0.84805	AdjIust2C	0	0.829	0.044	0.258	0.056	0.257	0.056	0.944	0.018	0.971	0.008	2.400	0.111	1.278	0.111	24.007	0.839	25.437	0.824
63	69111	IAcc	HistC	DIP	COR50	0.84805	AdjIust2C	0.5	0.825	0.046	0.267	0.058	0.220	0.050	0.942	0.018	0.972	0.008	1.756	0.088	1.268	0.106	23.655	0.913	25.671	0.904
64	69111	IAcc	HistC	noDIP	COR0	0.84805	20	1	0.780	0.056	0.324	0.076	0.208	0.058	0.909	0.035	0.966	0.010	1.121	0.000	1.144	0.097	20.482	0.996	23.409	0.367
65	69111	IAcc	HistC	noDIP	COR100	0.84805	20	0	0.801	0.051	0.295	0.066	0.293	0.067	0.924	0.029	0.965	0.009	2.309	0.103	1.188	0.103	21.551	0.802	23.243	0.353
66	69111	IAcc	HistC	noDIP	COR50	0.84805	20	0.5	0.792	0.054	0.308	0.069	0.256	0.062	0.919	0.031	0.965	0.009	1.708	0.081	1.169	0.103	21.078	0.860	23.295	0.384
67	69111	IAcc	LowMean	DIP	COR0	0.84805	AdjIust2C	1	0.828	0.041	0.273	0.069	0.204	0.055	0.947	0.017	0.970	0.008	1.992	0.193	1.259	0.095	23.416	0.620	25.123	0.654
68	69111	IAcc	LowMean	DIP	COR100	0.84805	AdjIust2C	0	0.834	0.039	0.264	0.067	0.264	0.067	0.950	0.015	0.969	0.009	3.264	0.222	1.272	0.101	23.732	0.438	24.811	0.485
69	69111	IAcc	LowMean	DIP	COR50	0.84805	AdjIust2C	0.5	0.832	0.040	0.267	0.066	0.237	0.059	0.950	0.016	0.970	0.009	2.629	0.207	1.268	0.103	23.594	0.480	24.913	0.554
70	69111	IAcc	LowMean	noDIP	COR0	0.84805	20	1	0.811	0.043	0.304	0.074	0.231	0.058	0.938	0.020	0.965	0.009	1.992	0.193	1.202	0.083	21.741	0.719	23.410	0.329
71	69111	IAcc	LowMean	noDIP	COR100	0.84805	20	0	0.818	0.041	0.291	0.068	0.291	0.068	0.942	0.018	0.965	0.009	3.208	0.223	1.217	0.088	22.132	0.579	23.279	0.325
72	69111	IAcc	LowMean	noDIP	COR50	0.84805	20	0.5	0.816	0.042	0.295	0.070	0.262	0.061	0.941	0.019	0.965	0.009	2.601	0.206	1.213	0.085	21.954	0.608	23.327	0.343
73	69111	IAcc	LowMean	DIP	COR0	0.84805	AdjIust2C	1	0.833	0.040	0.275	0.066	0.233	0.058	0.952	0.011	0.966	0.009	1.995	0.044	1.254	0.094	23.226	0.321	24.082	0.361
74	69111	IAcc	LowMean	DIP	COR100	0.84805	AdjIust2C	0	0.835	0.039	0.271	0.064	0.271	0.064	0.953	0.012	0.966	0.009	3.254	0.106	1.259	0.096	23.314	0.300	24.033	0.338
75	69111	IAcc	LowMean	DIP	COR50	0.84805	AdjIust2C	0.5	0.834	0.040	0.273	0.064	0.251	0.058	0.952	0.011	0.966	0.009	2.624	0.079	1.255	0.093	23.272	0.320	24.060	0.351
76	69111	IAcc	LowMean	noDIP	COR0	0.84805	20	1	0.825	0.042	0.289	0.067	0.245	0.061	0.950	0.012	0.963	0.010	1.995	0.044	1.217	0.087	22.453	0.462	23.294	0.342
77	69111	IAcc	LowMean	noDIP	COR100	0.84805	20	0	0.828	0.040	0.287	0.066	0.287	0.066	0.950	0.012	0.963	0.009	3.221	0.102	1.226	0.091	22.558	0.432	23.267	0.338
78	69111	IAcc	LowMean	noDIP	COR50	0.84805	20	0.5	0.826	0.041	0.289	0.067	0.266	0.061	0.950	0.012	0.963	0.009	2.606	0.075	1.221	0.089	22.503	0.448	23.283	0.349
79	69111	IAcc	RedVarC	DIP	COR0	0.84805	AdjIust2C	1	0.827	0.044	0.267	0.062	0.192	0.054	0.946	0.017	0.971	0.008	0.996	0.032	1.260	0.090	23.255	0.416	25.006	0.541
80	69111	IAcc	RedVarC	DIP	COR100	0.84805	AdjIust2C	0	0.832	0.041	0.260	0.063	0.260	0.063	0.948	0.016	0.970	0.008	2.269	0.091	1.273	0.091	23.565	0.380	24.797	0.504
81	69111	IAcc	RedVarC	DIP	COR50	0.84805	AdjIust2C	0.5	0.829	0.042	0.262	0.062	0.227	0.057	0.947	0.017	0.970	0.008	1.630	0.074	1.266	0.092	23.388	0.394	24.886	0.510
82	69111	IAcc	RedVarC	noDIP	COR0	0.84805	20	1	0.809	0.046	0.302	0.070	0.225	0.062	0.939	0.019	0.967	0.009	0.996	0.032	1.198	0.085	21.700	0.629	23.406	0.310
83	69111	IAcc	RedVarC	noDIP	COR100	0.84805	20	0	0.815	0.044	0.291	0.066	0.291	0.067	0.941	0.019	0.965	0.009	2.203	0.087	1.207	0.083	22.052	0.575	23.304	0.331
84	69111	IAcc	RedVarC	noDIP	COR50	0.84805	20	0.5	0.812	0.046	0.292	0.065	0.256	0.059	0.940	0.019	0.966	0.009	1.600	0.072	1.203	0.083	21.906	0.580	23.360	0.316