Decision Support Systems 80 (2015) 72–82

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Mathematical modeling and Bayesian estimation for error-prone retail shelf audits

Howard Hao-Chun Chuang *

College of Commerce, National Chengchi University, Taipei 11605, Taiwan

ARTICLE INFO

Article history:
Received 5 August 2014
Received in revised form 5 October 2015
Accepted 5 October 2015
Available online 22 October 2015

Keywords:
Retail operations
Audit services
Inspection error
Risk aversion
Bayesian inference

ABSTRACT

Prevalent execution errors such as out-of-stock, inventory record inaccuracy, and product misplacement jeopardize retail performance by causing low on-shelf availability, which discourages not only retailers who have lost sales but also manufacturers who have worked hard to deliver goods into retail stores. Thus, external service companies are hired by manufacturers to conduct manual inspection regularly. Motivated by the practical need of shelf audit service providers, we use a general cost structure to develop a decision support model for periodic inspection. Some qualitative insights about the intricate relationships among inspection efficacy, cost factors, failure rate of shelf inventory integrity, and optimal decisions are derived from analytics assuming risk-neutrality. From simulation experiments we also find that managers' risk preferences have non-trivial impacts on optimal decisions. Based on a total cost standpoint high-quality inspection is predominantly preferred regardless of the level of risk aversion. Finally, we propose a Bayesian statistical model and a Markov chain Monte Carlo approach to estimate model parameters such that managers can make empirically informed decisions. Our major contribution lies in developing a mathematical model that is practically applicable and proposing a Bayesian estimation approach to rationalize unobservable model parameters, which are influential to optimal decisions but often arbitrarily assumed by decision makers.

1. Introduction

Retail operations is composed of various tasks pertaining to assortment planning, product pricing, inventory optimization, and store execution [20]. Among those tasks, store execution is highly labor-extensive and complicated because it involves people, processes, and technology. Thus, execution errors such as shelf out-of-stock (OOS), inventory record inaccuracy (IRI), and product misplacement have become norms rather than anomalies even at financially successful retailers [47]. Store execution errors jeopardize retail performance by resulting in low on-shelf availability, which discourages not only retailers who have lost sales but also other supply chain members who have worked hard to deliver goods into the retail outlet. Being well-known for its operational excellence, Walmart recently admitted to a low on-shelf availability issue and predicted a $3 billion opportunity in filling in empty shelves [13].

Facing prevalent issues pertaining to on-shelf availability, retailers have gradually seen the need of allocating extra labor capacity to carry out shelf audits in order to reach higher service levels [17]. However, hiring more employees who are able to execute prescribed tasks and fix shelf errors goes against the common practice in retailing to minimize labor cost [22,46]. Since low on-shelf availability is a serious problem for retailers as well as *manufacturers* [41], manufacturers search for alternative approaches (rather than retailers' regular operations) to maximize the availability of their products [6,15].

A potential answer for manufacturers to fix low on-shelf availability is to ask external companies who provide shelf audit services to correct faulty items that may experience OOS, IRI, or misplacement [10]. Those external service agents are capable of working with different store formats (e.g., grocery, club, drug, convenience). Their associates excel in reshelving or display maintenance to complement retailers' regular operations, and conduct other tasks such as placing promotional goods. Moreover, periodic shelf audits performed by those thirty party companies are appealing to manufacturers because they also solve the conflict of retailers' potential bias to selectively report good audit outcomes [18]. Chuang et al. [10] report a successful case in which they conduct a field experiment in a U.S. retail chain and show that external audit services is a cost-effective way for product manufacturers to improve on-shelf availability.

Even though external shelf audits seem to be a promising solution to the recurring problem of low on-shelf availability, designing a cost optimal inspection policy for those external service companies turns out to be difficult because of two issues. First, information regarding inventory transactions may not be available to the service companies who have limited/no access to point-of-sale (POS) data. Second, it is nearly impossible to achieve error-free shelf inspection because an ordinary associate usually has to audit multiple items at multiple stores within a limited time. As a result, a certain amount of inspection error is inevitable and needs to be considered by decision makers. In attempt to tackle the

* Tel.: +886 2 29393091x88147; fax: +886 2 29393754.
E-mail address: chuang@nccu.edu.tw.

http://dx.doi.org/10.1016/j.dss.2015.10.003
0167-9236/

aforementioned issues, this paper presents a periodic inspection policy that triggers physical audits to increase on-shelf availability. We present a normative analysis of inspection decisions while taking into account inspectors' fallibility and managers' risk attitudes. Our paper addresses the question: *for external service providers with limited information about on-shelf items' status, what is the optimal frequency of shelf audits provided a level of inspection error and risk aversion?* We answer the question by deriving static analytics under risk neutrality and performing simulation studies under risk aversion.

The notion of *inspection error* and *risk aversion* is critical to our inspection policy design and makes our modeling effort relevant. On the one hand, as opposed to the commonly assumed "perfect inspection," we posit that any inspection in the real world can hardly be error-free. The reality is that inspection errors vary with human efforts and significantly increase the level of complexity surrounding the design of inspection policies [30]. Since the competencies, experiences, and motivations of inspectors are different, the probability of making mistakes will differ [5]. However, studies on the impact of error-prone inspection are scant in the context of retail shelf audits. We fill in the gap by explicitly incorporating human fallibility into our model and assessing the impact of different levels of inspection error rates.

On the other hand, numerous studies on inspection policies assume risk neutrality, which is valid only if optimal decisions are invariant with managers' risk attitudes [37]. Unfortunately, most of the earlier attempts (e.g., [28,36]) to optimize inspection decisions have not taken into account managers' risk preferences. Peecher et al. [42] point out that audit initiatives are by no means risk-free and there are different elements of risk – internal risk, control risk, and detection risk – surrounding inspection policy design. Those elements of risk lead to uncertainties in total cost of shelf audit efforts. Seeing that optimal decisions will depend on the degree of risk aversion, we take a utility-based approach to analyze how risk aversion affects the design of inspection policies. The principle of maximizing expected utility has a rich theoretical foundation [11] that enables us to explore the interaction between risk preferences and optimal decisions.

Our study makes several contributions. First, our model has a fairly general cost structure and it is built upon realistic assumptions of inspection efficacy and managerial risk preferences. Managers can adopt the proposed model to achieve cost-effective inspection and recover profit loss caused by low on-shelf availability. Our modeling effort is particularly relevant for retail service providers who need to periodically send associates into retail stores to maintain shelf inventory integrity. Second, our model considers imperfect inspection and accommodates two types of errors – the error of failing to correct faulty items and the error of miscorrecting accurate ones. Further, we assess managers' risk preferences that are found to have substantive impacts on optimal decisions. We observe that from a cost standpoint high-quality inspection (i.e., low error probability) is generally preferred regardless of the degree of risk aversion. Third, our model also captures the random degradation of on-shelf availability due to store execution errors. We find interesting dynamics among inspection efficacy, failure rate of shelf inventory, and cost factors. Our analysis shows that the ignorance of imperfect inspection and random shelf error generation would result in suboptimal audit decisions. Lastly, early papers make hypothetical assumptions about the distribution of inspection error [5,14] because no observable data can be applied to directly estimate the error distribution in a non-experimental context. We address the issue by making Bayesian inference about the level of inspection error instead of making hypothetical guesses. We adopt Bayesian hierarchical modeling and use a Metropolis-within-Gibbs sampling scheme to statistically infer unobservable human errors given observed inspection outcomes. To the best of our knowledge, there is no similar attempt reported in the literature of shelf inspection and inventory audits.

The rest of this article is organized as follows. Section 2 summarizes the relevant literature related to our work; the formulation and analysis of a periodic inspection model for shelf audit service providers under risk-neutrality and risk-aversion are presented in section 3 and section 4 respectively. In section 5 we propose a Bayesian methodology to estimate unknown model parameters that are crucial for optimal decisions. We conclude by articulating practical implications and research limitations.

2. Related literature

A stream of literature has engaged in developing decision support models for retail shelf audits. One of the seminal studies is by Hughes [28] who formulates a Markov decision process to determine the optimal timing of audits while considering the efficacy of auditing. Morey and Dittman [36] further propose a model to calculate the optimal timing of stock audits based on pre-specified goals of inventory accuracy. More recently, Sandoh and Shimamoto [45] devise a stochastic model to find the optimal frequency of inventory counting in a supermarket. Kok and Shang [33] propose a joint inventory inspection and replenishment policy that is capable of recovering a large proportion of benefits brought by RFID adoption. DeHoratius et al. [12] develop a shelf inspection policy based on expected value of perfect information. Atali et al. [3] also work on the problem of inventory integrity within periodic review inventory systems. Our model differs from previous studies in two major aspects. First, neither sales quantity nor inventory position is known to decision makers (i.e., managers of external service firms) who typically have limited observations on on-shelf items from periodic inspection. Second, we explicitly incorporate inspection efficacy and risk preferences into inspection policy design.

Extant studies on inspection assume risk neutrality, an assumption that is not likely to be valid in our context of retail shelf audits. Peecher et al. [42] define audit risk as the product of three underlying risks: inherent risk, control risk, and detection risk. Here *inherent risk* refers to the fact that on-shelf availability could easily be compromised due to various execution errors, which are likely to persist without internal controls [43]. However, imposing internal controls (e.g., periodic inspection) has *control risk* that is related to two cost factors – a cost of inspecting/correcting faulty items and a cost of leaving faulty items unfixed. Thus, control risk involves optimizing inspection decisions to minimize the sum of those costs. Lastly, *detection risk* refers to the fact that human inspectors are not able to detect and fix all errors. More often than not, inspectors contaminate inventory data as "large errors often remain in the stock records because of inaccuracies in the counting procedure" [29].

The three types of risks found in retail shelf audits shed light on the need for incorporating risk aversion into decision support models. However, most of the models discussed above focus on mitigating *inherent* and *control* risks without explicitly examining *detection* risk. As opposed to the commonly assumed "perfect inspection" in retail operations research [33], we posit that any inspection in the real world can hardly be error-free. The reality is that inspection efficacy varies with human efforts and significantly increases the level of complexity surrounding the inspection policy design. Since the competencies, experiences, and motivations of individual auditor differ, the probability of their making inspection errors will differ [5]. The impact of error-prone inspection has been widely studied in a manufacturing environment [14,48]. That said, studies on the impact of auditor error are scant in the context of retailing. We fill in the gap by formally analyzing the costs and benefits of different levels of inspection efficacy.

Aside from the above-mentioned studies on designing cost-minimization inspection policies, our paper is related to studies that apply statistical process control (e.g., [21,25]) or acceptance sampling (e.g., [16]; [19]) approaches to improve inventory integrity. However, most of the statistical approaches require actual and/or recorded inventory levels that are not available in our setting. Moreover, with respect to unobservable inspection efficacy, early papers make hypothetical assumptions about the distribution of inspection error (e.g., [5,14]) because no observations can be used to estimate the distributions directly in a non-experimental context. We address this limitation by