Share and Share Alike? The Value of Pooling Demand Data Across Retail Locations

by

Sarah M. Colby

B.Sc., The University of Guelph, 1999

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Faculty of Commerce and Business Administration)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

July 30, 2002

© Sarah M. Colby, 2002
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Faculty of Commerce and Business Administration

The University Of British Columbia
Vancouver, Canada

Date July 30, 2002
Abstract

We consider the value of pooling demand data across multiple retail outlets in a multiple-period newsvendor problem. In particular, our setting consists of a single firm with two newsvendor retail locations. The retailers each face normally distributed, stationary demand for a product. The variance of demand is known, but the mean is unknown and estimated. This estimate is updated as consumer demand at the retail locations is observed. We contrast the situation in which a central agent pools the retailers' demand data when updating the demand forecast with that when the retailers act with local data only. In a numerical study, we investigate the effects of the following factors on the value of pooling: (1) the relative size of the retailers' markets; (2) whether the product is a high-profit or a low-profit product; (3) the number of periods in the horizon; and (4) the forecast updating method. Further, we study the effects of these factors under a number of different scenarios characterized by the a priori beliefs about mean demand. Specifically, these scenarios represent initial beliefs about mean demand which are either overconfident, underconfident or reasonable. In our numerical study we find that the effect of the relative size of the retailers' markets on the value of pooling is not interesting. The value of pooling is greater for low-profit products than for high-profit products. As expected, pooling is more valuable when the a priori beliefs about mean demand are overconfident or underconfident than when they are reasonable. The effect of the forecast updating method on the value of pooling depends on the initial beliefs about mean demand. Likewise, the effect of the horizon length depends on the forecast updating method, as well as the initial beliefs about mean demand.
Contents

Abstract ............................................................... ii
Contents ............................................................... iii
List of Tables ......................................................... v
List of Figures ......................................................... vi
Acknowledgements .................................................. vii

I Thesis ................................................................. 1

1 Introduction ......................................................... 2

2 Literature Review .................................................. 4

3 Modelling Framework .............................................. 8
  3.1 Modelling Consumer Demand ............................... 8
  3.2 Cost Parameters and the Profit Function ................. 9
  3.3 The Subjective and Predictive Distributions ............. 9
  3.4 The Ordering Policy ....................................... 9
  3.5 Forecast Updating .......................................... 10
    3.5.1 Moving average ........................................ 10
    3.5.2 Exponential smoothing ................................ 11
    3.5.3 Bayesian updating ..................................... 11

4 Numerical Study .................................................. 13
  4.1 Experimental Design ..................................... 13
  4.2 Computational Details ..................................... 15
    4.2.1 Choosing the smoothing parameter and the order of the moving average ....... 15
    4.2.2 Generating uniform random numbers .................. 16
    4.2.3 Computing the percentage points of the standard normal distribution .... 16

5 Results .............................................................. 18
  5.1 Value of Pooling by Scenario ............................. 18
  5.2 Effect of the Relative Size of the Retailers' Markets ... 19
  5.3 Effect of the Critical Fractile ............................ 21
  5.4 Effect of the Horizon Length ............................. 23
  5.5 Effect of the Forecast Updating Method .................. 23
  5.6 Effect of the Horizon Length by Forecast Updating Method .... 28

6 Conclusions .......................................................... 33
Contents

Bibliography ......................................................... 35

A Sampling Distribution of the Moving Average Forecast ................. 37

B Sampling Distribution of the Exponential Smoothing Forecast ........... 38

C Algorithm AS 241 .................................................. 39

D Confidence Interval for a Median ..................................... 42
List of Tables

4.1 Definition of the scenarios characterizing the a priori beliefs about mean demand 14
4.2 Value of the smoothing parameter and the order of the moving average by scenario 16

5.1 Five-number summary of the median percentage improvements in profit with pooling by scenario 19
5.2 Five-number summary of the median percentage improvements in profit with pooling by scenario for instances with horizon lengths of 50 or less 19
5.3 95% C.I.'s for the median percentage improvement in profit with pooling by relative size of the retailers' markets and scenario 19
5.4 95% C.I.'s for the median percentage improvement in profit with pooling by critical fractile and scenario 21
5.5 95% C.I.'s for the median percentage improvement in profit with pooling by horizon length and scenario 25
5.6 95% C.I.'s for the median percentage improvement in profit with pooling by forecast updating method and scenario 27
5.7 Relative efficiency of the pooled versus the non-pooled exponential smoothing forecast in a sample of periods for scenarios I, II, and V 32
List of Figures

5.1 Effect of the relative size of the retailers' markets on the median percentage improvement in profit with pooling by scenario ...................... 20
5.2 Effect of the critical fractile on the median percentage improvement in profit with pooling by scenario ........................................ 22
5.3 Effect of the horizon length on the median percentage improvement in profit with pooling by scenario ........................................ 24
5.4 Effect of the forecast updating method on the median percentage improvement in profit with pooling by scenario ........................................ 27
5.5 Effect of the horizon length on the median percentage improvement in profit with pooling by scenario when a moving average forecast update is used .... 28
5.6 Effect of the horizon length on the median percentage improvement in profit with pooling by scenario when an exponential smoothing forecast update is used . 29
5.7 Effect of the horizon length on the median percentage improvement in profit with pooling by scenario when a Bayesian forecast update is used ............... 30
Acknowledgements

I would like to thank Dr. David Glenn and Dr. Martin L. Puterman for their helpful comments and guidance. I have learned a great deal about the nature of academic research and acquired many skills that will be invaluable as I continue my studies. Sincere thanks are also extended to Dr. Jonathan Berkowitz for his technical advice and encouragement. I greatly appreciated your words of wisdom throughout this process.

In addition, I would like to acknowledge the Faculty of Commerce and Business Administration, the Faculty of Graduate Studies, and the Natural Sciences and Engineering Research Council of Canada (NSERC) for their financial support. I would also like to acknowledge the Centre for Operations Excellence (COE) for the use of its resources. Finally, I would like to express heartfelt gratitude to my parents—Alice and Ross Colby—and to Dao Le for their immeasurable patience and support.
Part I

Thesis
Chapter 1

Introduction

The topic of supply chain management is at the forefront of much discussion in the business press today. Many companies have realized that applying a “total systems approach” to managing the flow of information, goods, and services—from procuring raw materials to satisfying the end customer—is crucial to their success [5, p. 330]. Consequently, a great deal of effort is being invested in improving supply chain efficiency.

From a historical perspective, the focus of supply chain management has been on the physical flow of goods through a value-adding chain. An important area in which a company could gain a competitive advantage was transportation and logistics. More recently, however, the concept of supply chain management has broadened to include managing the flow of information through the supply chain. This view emphasizes the need for collaboration between the supply chain members. The internet affords such cooperation, and, in light of this, advances in information technology have had a tremendous impact on supply chain management. Scanners, for example, collect point-of-sales (POS) data, and these data can be shared throughout the supply chain using electronic data interchange (EDI). Such developments have lead to a belief among management that sharing information is key to improving supply chain performance, but is this really the case [4]? Intuitively, information is beneficial, but are there conditions or factors which affect the magnitude of these benefits?

There is a growing interest in these questions in both the business and academic communities. The types of information shared are numerous, but the most common are inventory levels, sales data, and demand forecasts [14]. Sharing inventory information, for instance, can decrease the total inventory in the supply chain. Centrally coordinated inventory management can help prevent holding multiple safety inventories or having stockouts at multiple locations. Applications of sharing inventory data include Continuous Replenishment Programs (CRP) and Vendor-Managed Inventory (VMI). Wal-Mart, for example, shares with Proctor & Gamble the inventory information regarding their products because Proctor & Gamble can manage their own products better than can Wal-Mart [14].

An often cited motivation for sharing sales data is the phenomenon of increasing demand variability as one moves upstream in a supply chain—the well-known “bullwhip effect”. This distortion of demand results in suppliers poorly forecasting demand, carrying excess inventory, and providing poor customer service [14]. These effects can be mitigated if retailers share their POS data with their suppliers. Wal-Mart’s Retail Link program is an example of this, wherein suppliers such as Johnson & Johnson and Lever Brothers have on-line access to Wal-Mart’s sales data [13].

Lastly, sharing demand forecasts allows a retailer to benefit from a supplier’s expert knowledge. In Wal-Mart’s Collaborative Forecasting and Replenishment initiative (CFAR), suppliers such as the pharmaceutical manufacturer Warner-Lambert and Wal-Mart jointly develop demand forecasts, taking advantage of the fact that Warner-Lambert, for instance, knows how weather conditions are likely to affect the sales of its products [14].

The examples above involve information sharing between a retailer and an upstream supplier. In contrast, we consider information sharing between multiple retail outlets. Specifically, we study the value in pooling demand data across retail locations in a setting involving a single firm and two newsvendor retailers. The retailers each face normally distributed demand whose
Chapter 1. Introduction

variance is known but whose mean is unknown. Mean demand is estimated, and this estimate is updated as consumer demand at the retail locations is observed.

We contrast the situation in which a central agent pools the retailers’ demand data when forecasting demand with that when the retailers act with local data only. In a numerical study, we investigate the effects of the following factors on the value of pooling: (1) the relative size of the retailers’ markets; (2) whether the product is a low-profit or a high-profit product; (3) the number of periods in the horizon; and (4) the forecast updating method. We consider three forecast updating methods based on a moving average, exponential smoothing and Bayes' Theorem respectively. The moving average forecast updating method is of interest because it reflects a simple forecasting technique which is commonly used in practice. The exponential smoothing forecast updating method is included because it is also commonly used but is regarded as more sophisticated than a moving average. Lastly, the Bayesian forecast updating method is studied because it is a theoretical framework often modelled in the supply chain literature.

Furthermore, we study the effects of these factors under a number of different scenarios characterized by the a priori beliefs about mean demand. In particular, we consider when these beliefs are overconfident (i.e., the central agent and the retailers are confident in incorrect beliefs), underconfident (i.e., they lack confidence in correct beliefs) or reasonable (i.e., their degree of confidence is consistent with how correct their beliefs are).

Some of the results of our numerical study are as follows: The relative size of the retailers’ markets does not have an interesting effect on the value of pooling. The value of pooling is greater for low-profit products than for high-profit products. As expected, pooling is more valuable when the a priori beliefs are overconfident or underconfident than when they are reasonable. Regardless of the initial beliefs about mean demand, however, there is more value in pooling when a moving average forecast update is used than when an exponential smoothing forecast update is used. When the a priori beliefs are overconfident, pooling is most valuable when a Bayesian forecast update is used and least valuable in the case of an exponential smoothing forecast update. When the a priori beliefs are underconfident, on the other hand, pooling is most beneficial when a moving average forecast update is used and least beneficial in case of a Bayesian forecast update. The effect of the horizon length on the value of pooling depends on the initial beliefs about mean demand as well as the forecast updating method.

The remainder of this paper is organized as follows. In Chapter 2 we overview the existing academic literature in the area of information sharing. The details of our modelling framework are addressed in Chapter 3. Chapter 4 describes the numerical study, providing further computational details regarding the simulation and outlining the experimental design. The results are presented in Chapter 5. Finally, in Chapter 6 we conclude and highlight areas for future research.
Chapter 2

Literature Review

Many companies have invested heavily in information technology as a means of better managing and reducing inefficiency in their supply chains. This investment "has resulted in the availability of more information on channel activities to decision maker[s], who, in turn, must find ways of incorporating [it into] their day-to-day decisions" [16, §1.1]. There is a widely held belief that capturing and sharing this information will improve supply chain performance. Is this the case, and, if so, are the benefits significant? Moreover, what conditions affect the magnitude of these benefits? The literature which attempts to answer these questions is relatively recent, and the conclusions differ considerably from paper to paper. In this chapter, we present an overview of this research. In particular, the following papers study information sharing between one or more retailers in a supply chain and their upstream suppliers.

Lee et al. [13] use analytical models to study the value of information sharing in a two-level supply chain with a single retailer and a single manufacturer. Demand is a simple autocorrelated AR(1) process. When there is no information sharing, the manufacturer knows only the retailer's order quantity; with information sharing, the manufacturer also knows demand at the retailer. The manufacturer benefits from the additional information because it allows him or her to better forecast the retailer's future orders. They compare the manufacturer's order quantities—using an \((s, S)\) order-up-to policy—with and without information sharing. The benefit of information sharing is measured in terms of the reduction in the manufacturer's expected inventory holding and shortage costs and the reduction in the manufacturer's average on-hand inventory. In numerical examples, they find that the value of information sharing is the greatest when demand is highly autocorrelated, highly variable, or when the retailer's replenishment leadtime is long.

Raghunathan [17], however, argues that the benefits of information sharing reported by Lee et al. are an artifact of an assumption they made: When there is no information sharing, the manufacturer forecasts the retailer's future orders using only the most recent retailer order rather than the retailer's complete order history. Using analytical methods as well as a numerical study, Raghunathan shows that when this additional, available information is used—and when the demand process is known by both the retailer and the manufacturer, as Lee et al. assume—information sharing is of little value. Even if the retailer does not share his or her demand information with the manufacturer, the manufacturer can fairly accurately infer it from the retailer's order history and the parameters of the demand process. He or she can then use this to better forecast the retailer's order quantities. Moreover, the value of information sharing decreases with time because the manufacturer accumulates more and more order information. Using the model proposed by Lee et al., Raghunathan finds the maximum percentage decreases in cost (when the autocorrelation is highest, the demand variability is highest, and the retailer's replenishment leadtime is longest) are 18.65%, 17.65% and 13.26% respectively. When the manufacturer forecasts the retailer's future orders using the retailer's complete order history, on the other hand, Raghunathan finds that the percentage decreases in cost reported above are all essentially 0%.

Cachon and Fisher [4] study the value of sharing inventory data in a setting with one supplier and \(N\) identical retailers. The demand distribution is known and stationary. They compare supply chain costs using a traditional information policy (in which orders are the only information the retailers share with their supplier), with a full information policy (in which the retailers also
share their inventory data). An \((R, Q)\) reorder point/order quantity policy is used by both the supplier and the retailers in the case of the traditional information policy. In the case of the full information policy, the retailers use reorder point/order quantity policies but the supplier does not. Rather, he or she takes advantage of the additional information to improve ordering decisions and the allocation of inventory between the retailers. In a numerical study, Cachon and Fisher find that supply chain costs are 2.2% lower on average using the full information policy compared to the traditional information policy. The maximum difference they report is 12.1%. They contrast the value of information sharing with other benefits of information technology, namely, reduction in time and cost to process orders. Faster order processing results in shorter lead times; cheaper order processing leads to smaller batch sizes. In their numerical study, they find that cutting lead times in half and cutting batch sizes in half reduces supply chain costs by 21% and 22% on average respectively. Ultimately, they conclude that using information technology as a means of more efficiently moving goods through the supply chain is of substantially more benefit than using information technology to facilitate information sharing.

Gavirneni et al. [8] study the value of information in a capacitated supply chain with a single retailer and a single supplier. The retailer uses an order-up-to policy. In particular, they consider three models: a baseline case of no information sharing, a partial information sharing case, and a complete information sharing case. In the baseline setting, the supplier observes only the retailer's orders and knows neither the demand process nor the retailer's ordering policy. As Gavirneni et al. remark, this model is artificial since the supplier could, if he or she had reason to believe an order-up-to policy was being used, infer the parameters of this policy from the past demand data. Nevertheless, Gavirneni et al. assume the supplier does not do this. In the partial information setting, the supplier knows the demand process, as well as the fact that the retailer uses an order-up-to policy with the reorder point \(s\) and the order-up-to point \(S\) known. Based on this information (and the number of periods since the retailer last placed an order), the supplier can determine the probability of the retailer placing an order in a period and the distribution of the size of the order. In the complete information setting, the supplier also observes the supplier's demands each period. As in the partial information setting, the supplier can determine the probability of the retailer placing an order in a period and the distribution of the size of the order. In a numerical study, Gavirneni et al. investigate the cost savings to the supplier with information sharing. Specifically, they consider the effect of the supplier's capacity, the ratio of the supplier's shortage penalty to holding cost, the variance of demand, and the parameters of the retailer's order-up-to policy. In each instance, the cost associated with the no information setting is the greatest, and the cost of the complete information setting is the least. The percentage savings comparing the partial information setting to the no information setting range from 10% to 90%. As mentioned earlier, this is an artificial comparison. The percentage savings comparing the complete information setting to the partial information setting range from 1% to 35%. The percentage savings are the greatest for moderate demand variances, moderate differences \(S - s\), and high capacities.

Moinzadeh [16] studies the benefits of information sharing in a supply chain with one supplier and multiple retailers. Demand at the retailers is stationary and Poisson, and they use a reorder point/order quantity policy. He assumes that the supplier has access to the retailers' inventory levels and uses this information when making his or her replenishment decisions. In particular, when a retailer's inventory reaches a certain level (above or equal to the reorder point \(R\)), the supplier places an order of size \(Q\) from his or her source in anticipation of that retailer's future order. Acting in this proactive fashion, the supplier is better able to satisfy the retailers' demands, lowering system costs. Moinzadeh derives an analytical expression for the expected total cost of the system per unit time (average total cost rate) but uses numerical approximation to compare the cases of information sharing and no information sharing. The savings reported are 3.2% on average with a maximum savings of 34.9%. Information sharing is most beneficial when the supplier has a long lead time, the number of retailers is not large, when the ratio of
Chapter 2. Literature Review

The retailers' holding cost to that of the supplier is moderate, and when the batch size $Q$ is moderate.

Chen [6] studies a serial inventory system consisting of $N$ stages, wherein stage 1 orders from stage 2, stage 2 orders from stage 3, and so on. Stage $N$ orders from an outside source. In particular, he considers two types of ordering policies: In the first—an echelon-stock policy—an order is placed whenever the inventory position of the subsystem consisting of a stage and all its downstream stages falls below an echelon reorder point; in the second—an installation-stock policy—an order is placed whenever the local inventory position of a stage falls below an installation reorder point. The echelon-stock policy requires that each stage know the demand information at stage 1. That is, it requires centralized demand information. Chen develops an algorithm for determining the optimal echelon reorder point for each stage and a heuristic algorithm for determining the installation reorder points. The objective is to minimize long-run average total cost. In a numerical study, he examines the value of centralized demand information as measured by the relative cost difference between the echelon-stock policy and the installation-stock policy. The value of information reported is 1.75% on average with a maximum of 9%. Greater values of information are realized when there are a large number of stages in the system, when the lead times are long, or when the batch sizes are large. When demand variability is high, or the desired level of customer service is extreme, the value of information is less.

Aviv [1] studies the effect of two methods of forecast updating on supply chain costs. Specifically, he considers a setting with a single retailer and a single supplier. Inventory at the retailer and the supplier are replenished periodically using order-up-to policies based on their forecasts for future demand. The retailer and the supplier cooperate when setting their replenishment policies to minimize long-run average supply chain costs. In the local forecast model (LF), the variability of demand in a period is a combination of the variances of a series of "adjustment" terms—these reflect forecast updates made separately by the retailer and the supplier at the start of each period—and any residual uncertainty that remains. In this way, the supplier and the retailer can reduce the variance of their forecast errors. That is, over time, they are able to resolve demand uncertainty to some degree. Aviv allows the retailer's and the supplier's adjustments in a period to be correlated since their local adjustments may be based on the same information.

In the collaborative forecast model (CF), the retailer and the supplier share information about future demand and make joint forecast adjustments. Specifically, the retailer and the supplier each produce local adjustments (as in the LF setting) and, using a weighted average of these local adjustments (where the weights depend on their "forecasting capabilities" and the correlation between their local adjustments), a collaborative adjustment is determined. Aviv shows that the reduction in the variance of the forecast error from one period to the next is greater in the CF setting than in the LF setting. He also compares the LF and CF models to a baseline model (B) in which forecast adjustments are not incorporated into the replenishment policy.

A simulation-based approach is taken to determine whether, and by how much, supply chain costs decrease when collaborative forecasting is used. In particular, three comparisons are considered: (1) LF versus B; (2) CF versus B; and (3) CF versus LF. Aviv concludes that there are substantial benefits associated with implementing either CF or LF compared to B. He reports average decreases in supply chain costs of 11.1% and 19.4% respectively. The average decrease in supply chain costs associated with moving from a local forecasting practice to a collaborative forecasting practice is found to be 9.6%. The benefits of local forecasting are found to depend mainly on the measures of the retailer's and the supplier's forecasting capabilities. The additional benefits of implementing collaborative forecasting are found to depend mainly on the correlation between the retailer's and the supplier's forecast adjustments. Intuitively, if the retailer and the supplier each offer different information, there is more value in sharing their
information. Finally, Aviv also remarks that the benefits of collaborative forecasting are greater when the supplier's lead time is short. If the retailer shares information with the supplier, but the supplier's long lead time prevents him or her from using it, there is little or no benefit to this information.

The papers discussed in this chapter study the value of sharing various types of information in a number of inventory systems and using several ordering policies. The reported results vary considerably, and the differences in modelling approaches make direct comparison of these studies difficult. Raghunathan, however, concludes with an elegant insight that is worth keeping in mind with respect to our research: "for information sharing to be useful, the information shared should not be inferential by the receiving party using any of the available data" [17].

We are also interested in the value of information sharing, although we study this in a multiple-period newsvendor setting rather than a periodic review inventory system as described in the papers above. Also, whereas these papers analyze the benefit of information sharing between downstream and upstream supply chain members, we study the value of information sharing between retailers. Specifically, we consider the value to a single firm of pooling demand data across multiple retail locations. We measure the value of pooling as the percentage improvement in the firm's profit when the retailers' data are pooled. Our modelling framework is presented in greater detail in the next chapter.
Chapter 3

Modelling Framework

In this chapter we formulate our model. Specifically, we consider a single firm with two retail locations each facing a newsvendor ordering problem. There is also a central agent who is responsible for choosing the retailers' order quantities. In each period—over a finite horizon of $l$ periods—the sequence of events is as follows: (1) the central agent places an order for each retail location, and these orders are immediately received; (2) the retailers observe consumer demand and share this information with the central agent; and (3) using the pooled retailer demand information, the central agent estimates mean demand and constructs a predictive distribution for demand at each retail location. We compute the average profit over the $l$ periods and compare this to the case when the retailers act with local data only. Sections 3.1 through 3.5 specify the model for consumer demand; give the cost parameters and the profit function; define the subjective and predictive distributions; detail how observed demand is used to update the parameters of the predictive distribution; and state the ordering policy.

3.1 Modelling Consumer Demand

We let $D_{t,1}$ and $D_{t,2}$, for $t = 1, 2, \ldots, l$, denote consumer demand for the product in period $t$ from retailer 1 and retailer 2 respectively. We assume independence of demand between the retailers and across the periods. There is a known relationship between demand from the two retail locations wherein retailer 2 operates in a market which is a known constant $k$ times that of retailer 1. Specifically, for $t = 1, 2, \ldots, l$,

$$D_{t,1} \sim N(\theta, \sigma^2),$$

and

$$D_{t,2} \sim N(k\theta, k^2\sigma^2).$$

The variance of demand $\sigma^2$ known, but mean demand $\theta$ unknown.

For notational convenience, we let

$$X_{t,1} \equiv D_{t,1},$$

and

$$X_{t,2} \equiv D_{t,2}/k.$$

We let $x_{t,1}$ and $x_{t,2}$ be the respective realizations of $X_{t,1}$ and $X_{t,2}$. Hence, $X_{t,1}$ and $X_{t,2}$, for $t = 1, 2, \ldots, l$, are independent and identically distributed (i.i.d.) $N(\theta, \sigma^2)$.

We assume that the retailers’ demands are not censored by their order quantities: Real demand during the period is observed. In the case of censored demand, lost sales are unobserved. That is, if a retailer’s order quantity exceeds demand in a period, demand for the product is observed; if a retailer orders fewer units than demand, only the number of units sold is observed. Censored demand is beyond the scope of this paper.
3.2 Cost Parameters and the Profit Function

At the beginning of each period, the retailers purchase the product from a supplier for unit wholesale price $w$. Throughout the period, they sell to the consumer for unit selling price $p$. At the end of the period, any unsold units are salvaged for unit salvage value $s$. We assume $p > w > s$. Note that our model does not include penalty costs for unsatisfied demand. Also, we do not allow backorders. Given an order quantity $q_{t,i}$ and observed demand $d_{t,i}$, the profit realized by retailer $i$, $i = 1, 2$, in period $t$ is

$$(p - w)d_{t,i} - ((p - w)(d_{t,i} - q_{t,i})^+ + (w - s)(q_{t,i} - d_{t,i})^+),$$

where $p - w$ is the underage cost (marginal profit), $w - s$ is the overage cost (marginal loss), $(d_{t,i} - q_{t,i})^+$ is the number of units of unsatisfied demand, and $(q_{t,i} - d_{t,i})^+$ is the number of excess units ordered.\(^1\) In other words, a retailer's profit is the profit which would have been realized if his or her order quantity matched demand, less the profit lost from ordering either too few or too many units. The profit to the firm is the sum of the retailers' profits. That is,

$$\sum_{i=1,2} (p - w)d_{t,i} - ((p - w)(d_{t,i} - q_{t,i})^+ + (w - s)(q_{t,i} - d_{t,i})^+).$$

3.3 The Subjective and Predictive Distributions

Recall from §3.1, $X_{t,1}$ and $X_{t,2}$, $t = 1, 2, \ldots, l$, are i.i.d. $N(\mu, \sigma^2)$ where the variance of demand $\sigma^2$ is known, yet mean demand $\mu$ is unknown. Prior to observing demand in the first period, the beliefs regarding $\mu$ are summarized by a subjective distribution. In particular, the central agent feels that $\mu \sim N(\mu', \sigma^2)$ with $\mu'$ and $\sigma^2$ known. After observing the retailers' demands each period, the central agent updates the estimate of mean demand—the demand forecast. Based on the forecast, he or she updates the predictive distribution for demand. Recall that we compare this to when the retailers act with local data only. In this case, the retailers share the central agent's beliefs about $\mu$.

In the following, we introduce notation which will be used in §3.5. We define $\hat{\theta}_{t,p}$ to be the forecast at the end of period $t$ based on the pooled retailer demand information. The pooled predictive distribution at the end of period $t$, $F_{t,p}$, is chosen to be $N(\hat{\theta}_{t,p}, V_{t,p} + \sigma^2)$, where $V_{t,p}$ is the variance of the pooled forecast. Note that the variance of the predictive distribution consists of two components: the first, attributable to the variance of the forecast; and the second, to the variance of demand. When the retailers act using local data only, each retailer has his or her own local forecast at the end of period $t$. We will denote these by $\hat{\theta}_{t,1}$ and $\hat{\theta}_{t,2}$ for retailer 1 and retailer 2 respectively. Likewise, they each have their own predictive distribution at the end of period $t$, $F_{t,i}$, where $F_{t,i}$ is $N(\hat{\theta}_{t,i}, V_t + \sigma^2)$ for $i = 1, 2$. The variance of the forecast $V_t$ is independent of the retailer $i$.

3.4 The Ordering Policy

In the familiar newsvendor model, the optimal order quantity (that which maximizes expected profit) depends on the probability distribution of demand and the critical fractile—the ratio of the underage cost (marginal profit) to the sum of the overage cost (marginal loss) and the underage cost. Specifically, the order quantity $q$ is the solution to

$$F(q) = c,$$

\(^1\)We use the notation $x^+$ to denote $\max(0, x)$. 
where \( F \) is the cumulative density function of the demand distribution and \( c \) is the critical fractile.

In our model, the retailers' order quantities are based on the predictive distribution(s). From §3.2, the underage cost and the overage cost are \( p - w \) and \( w - s \) respectively. Hence, the critical fractile is \( c = (p - w)/(p - s) \). We let \( q_{t+1,1,i} \) denote the order quantity for retailer \( i, i = 1, 2 \), at the start of period \( t + 1 \) when the central agent pools the retailers' demand information. Similarly, we let \( q_{t+1,i} \) denote the order quantity for retailer \( i, i = 1, 2 \), at the start of period \( t + 1 \) when the retailers act with local data only. That is,

\[
q_{t+1,1,i} = F_{t+1,1}^{-1}(c),
q_{t+1,2,i} = kF_{t+1,2}^{-1}(c),
q_{t+1,1} = F_{t+1,1}^{-1}(c),
q_{t+1,2} = kF_{t+1,2}^{-1}(c).
\]

### 3.5 Forecast Updating

We consider three methods of forecast updating. These are based on a moving average, exponential smoothing, and Bayes' Theorem respectively. In this section, for each forecast updating method, we give the demand forecast and its variance when the central agent pools the retailers' demand information, and we give the local forecasts when the retailers act with local data only and their common variance.

#### 3.5.1 Moving average

The pooled moving average forecast at the end of period \( t \), for \( t \geq N \), and its variance are

\[
\hat{\theta}_{t,p} = \frac{1}{N} \sum_{j=1}^{N} x_{t-j+1,p}
\]

and

\[
V_{t,p} = \frac{\sigma^2}{2N}
\]

respectively, where

\[
x_{t-j+1,p} = \frac{x_{t-j+1,1} + x_{t-j+1,2}}{2},
\]

and \( N \) is the order of the moving average. When the retailers act with local data only, their local moving average forecasts at the end of period \( t \), for \( t \geq N \), and the common variance of these forecasts are

\[
\hat{\theta}_{t,i} = \frac{1}{N} \sum_{j=1}^{N} x_{t-j+1,i}, \quad \text{for } i = 1, 2,
\]

and

\[
V_{t} = \frac{\sigma^2}{N}
\]

respectively.

The subjective distribution is used to initialize the forecast updating method. Specifically, we let

\[
\hat{\theta}_{0,p} \equiv \hat{\theta}_{0,1} \equiv \hat{\theta}_{0,2} \equiv \mu
\]

and

\[
V_{0,p} \equiv V_{0,1} \equiv V_{0,2} \equiv \tau^2.
\]
When too few periods of demand have been observed to compute a moving average of order $N$, a simple average is used. In particular, for $1 \leq t < N$,

$$\hat{\theta}_{t,p} = \frac{1}{t} \sum_{j=1}^{t} x_{t-j+1,p},$$

$$V_{t,p} = \frac{\sigma^2}{2t},$$

$$\hat{\theta}_{t,i} = \frac{1}{t} \sum_{j=1}^{t} x_{t-j+1,i},$$

and

$$V_t = \frac{\sigma^2}{t}.$$

### 3.5.2 Exponential smoothing

The pooled exponential smoothing forecast at the end of period $t$, for $t \geq 1$, and its variance are

$$\hat{\theta}_{t,p} = \alpha x_{t,p} + (1 - \alpha) \hat{\theta}_{t-1,p}$$

and

$$V_{t,p} = \frac{\alpha}{2 - \alpha} \left(1 - (1 - \alpha)^{2t}\right) \frac{\sigma^2}{2} + (1 - \alpha)^{2t} \tau^2$$

respectively, where

$$x_{t,p} = \frac{x_{t,1} + x_{t,2}}{2},$$

and $\alpha \in [0,1]$ is the smoothing parameter. Recall that $\tau^2$ is the variance of the subjective distribution discussed in §3.3. When the retailers act with local data only, their local exponential smoothing forecasts at the end of period $t$, for $t \geq 1$, and the common variance of these forecasts are

$$\hat{\theta}_{t,i} = \alpha x_{t,i} + (1 - \alpha) \hat{\theta}_{t-1,i}, \quad \text{for } i=1,2,$$

and

$$V_t = \frac{\alpha}{2 - \alpha} \left(1 - (1 - \alpha)^{2t}\right) \sigma^2 + (1 - \alpha)^{2t} \tau^2$$

respectively. The derivation of the variance of the exponential smoothing forecast is given in Appendix B. Again, we use the subjective distribution to initialize the forecast updating method. That is,

$$\hat{\theta}_{0,p} \equiv \hat{\theta}_{0,1} \equiv \hat{\theta}_{0,2} \equiv \mu$$

and

$$V_{0,p} \equiv V_{0,1} \equiv V_{0,2} \equiv \tau^2.$$

### 3.5.3 Bayesian updating

When Bayes' Theorem is used to update the demand forecast, the subjective distribution has a special meaning—it is the prior distribution of $\theta$ at the beginning of period 1. Upon observing demand in period 1, this prior distribution is updated, resulting in a posterior distribution of $\theta$ (conditional upon the observed demand). In subsequent periods, the prior distribution at the start of a period is the posterior distribution from the end of the previous period. In the case of a Normal prior and a Normal likelihood, the posterior distribution is also Normal. A more in-depth discussion can be found in [15].
When the central agent pools the retailers' demand data, the mean and variance of the posterior distribution of $\theta$ at the end of period $t$, for $t \geq 1$, are

$$
\mu_{t,p} = \mu_{t-1,p} \left( \frac{1}{(\tau_{t-1,p}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right) + x_{t,p} \left( \frac{\left(\frac{\sigma^2}{2}\right)^{-1}}{(\tau_{t-1,p}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right)
$$

and

$$
\tau_{t,p}^2 = \left(\frac{1}{(\tau_{t-1,p}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right)^{-1}
$$

respectively, where

$$
x_{t,p} = \frac{x_{t,1} + x_{t,2}}{2}.
$$

When the retailers act with local data only, the means and the common variance of their local posterior distributions of $\theta$ at the end of period $t$, for $t \geq 1$, are

$$
\mu_{t,i} = \mu_{t-1,i} \left( \frac{1}{(\tau_{t-1,i}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right) + x_{t,i} \left( \frac{\left(\frac{\sigma^2}{2}\right)^{-1}}{(\tau_{t-1,i}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right), \text{ for } i=1,2,
$$

and

$$
\tau_{t,i}^2 = \left(\frac{1}{(\tau_{t-1,i}^2)^{-1} + \left(\frac{\sigma^2}{2}\right)^{-1}} \right)^{-1}
$$

respectively.

The pooled Bayesian forecast at the end of period $t$, for $t \geq 0$, and its variance are

$$
\hat{\theta}_{t,p} = \mu_{t,p}
$$

and

$$
V_{t,p} = \tau_{t,p}^2
$$

respectively. When the retailers act with local data only, their local Bayesian forecasts at the end of period $t$, for $t \geq 0$, and the common variance of these forecasts are

$$
\hat{\theta}_{t,i} = \mu_{t,i}, \text{ for } i=1,2,
$$

and

$$
V_t = \tau_t^2
$$

respectively.

As implied earlier, we let

$$
\mu_{0,p} \equiv \mu_{0,1} \equiv \mu_{0,2} \equiv \mu
$$

and

$$
\tau_{0,p}^2 \equiv \tau_{0,1}^2 \equiv \tau_{0,2}^2 \equiv \tau^2.
$$

Thus,

$$
\hat{\theta}_{0,p} = \hat{\theta}_{0,1} = \hat{\theta}_{0,2} = \mu
$$

and

$$
V_{0,p} = V_{0,1} = V_{0,2} = \tau^2.
$$
Chapter 4

Numerical Study

In this chapter we detail the numerical study. The experimental design is described in §4.1. Namely, the values of the system parameters and the reasons for these choices are presented; the different scenarios characterizing the a priori beliefs about mean demand are defined; and the response measured is given. Also, we explain how common random numbers (CRN) is used as a variance reduction technique. Computational details which were not addressed in Chapter 3 are presented in §4.2. We justify our choice of the smoothing parameter in the case of an exponential smoothing forecast update and the choice of the order of the moving average in the case of a moving average forecast update. In addition, we discuss the random number generator used to implement the numerical study in C++ as well as the algorithm used for computing the percentage points of the standard normal distribution.

4.1 Experimental Design

The numerical study consists of the 378 instances resulting from all combinations of the following system parameters:

\[ c \in \{0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95\}; \quad p = 1 + c; \quad w = 1; \quad s = c; \]
\[ k \in \{1, 2, 5\}; \quad l \in \{2, 5, 10, 50, 500, 1000\}; \]

and the choice of the forecast updating method—based on either a moving average, exponential smoothing or Bayes’ Theorem. Recall that \( c \) denotes the critical fractile, \( p \) is the selling price, \( w \) is the wholesale price, \( s \) is the salvage value, \( k \) is the relative size of the retailers’ markets, and \( l \) is the horizon length. In all instances, the variance of demand is \( \sigma^2 = 16 \) and mean demand is \( \theta = 20 \). The variance of demand is chosen to be large enough to provide interesting results, yet small enough to ensure a low probability of a retailer observing a negative demand in a period. Given our choice, this probability is \( 2.87 \times 10^{-7} \), and, in fact, no negative demands occurred. The choice of the cost parameters \( p, w \) and \( s \) is discussed in Chapter 6.

The experimental conditions above represent a wide range of situations: when the critical fractile is extremely small compared to when it is extremely large; when the sizes of the retailers’ markets are identical and when they are not; and when the horizon length is very short versus when it is very long. Although, in practice, we do not expect to observe critical fractiles as low as 0.05, 0.1 or 0.25, nor as high as 0.95, these values are included to provide insight into the value of pooling when the optimal newsvendor order quantity falls in either the lower or the upper tail of the predictive distribution. Both very short and very long horizon lengths are included to study the value of pooling in the initial periods as well as its long-run behaviour. The values of interest of the relative size of the retailers’ markets \( k \) initially also included 0.2 and 0.5. These were ultimately excluded from the numerical study since a value of 0.2 (indicating retailer 2’s market is 20% of the size of retailer 1’s market) is equivalent to a value of 5 in the sense that, in both cases, one retailer’s market is 500% of the size of the other’s. Because we are assuming the central agent has a subjective distribution common to both retailers (and likewise, when the retailers act using local data only), there is no need to distinguish between the size of retailer 1’s market relative to that of retailer 2 and the size of retailer 2’s market relative to
Chapter 4. Numerical Study

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$\mu/\theta$</th>
<th>$\tau/\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td>II</td>
<td>1.15</td>
<td>0.1</td>
</tr>
<tr>
<td>III</td>
<td>0.5</td>
<td>0.02</td>
</tr>
<tr>
<td>IV</td>
<td>1.5</td>
<td>0.02</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>VI</td>
<td>1</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 4.1: Definition of the scenarios characterizing the a priori beliefs about mean demand

that of retailer 1. This would be of importance, however, if the model allowed the retailers to have local subjective distributions. This will be discussed further in Chapter 6.

Each of the 378 instances are evaluated using many values for the parameters of the subjective distribution $\mu$ and $\tau$. Specifically, we consider six scenarios (I - VI) as characterized by the accuracy of the subjective distribution—defined as $\mu/\theta$—and the precision of the subjective distribution—defined as $\tau/\mu$.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$\mu/\theta$</th>
<th>$\tau/\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.85</td>
<td>0.1</td>
</tr>
<tr>
<td>II</td>
<td>1.15</td>
<td>0.1</td>
</tr>
<tr>
<td>III</td>
<td>0.5</td>
<td>0.02</td>
</tr>
<tr>
<td>IV</td>
<td>1.5</td>
<td>0.02</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>VI</td>
<td>1</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 4.1 lists these scenarios. In scenarios I and II, for example, the subjective distribution is reasonably accurate and reasonably precise. In scenarios III and IV, on the other hand, the subjective distribution is very inaccurate yet highly precise. In such cases, we say the a priori beliefs about mean demand are "overconfident". In contrast, in scenario V, the subjective distribution is highly accurate but also highly imprecise. We say the a priori beliefs about mean demand are "underconfident". Lastly, in scenario VI, the subjective distribution is highly accurate and highly precise.

Two responses are measured for each scenario/instance: (1) the average profit over the horizon with pooling; and (2) the average profit over the horizon without pooling. The value of pooling is measured as the following percentage improvement:

$$\% \text{ improvement} = \frac{\text{with pooling} - \text{without pooling}}{|\text{without pooling}|} \times 100\%.$$  

Each scenario/instance is replicated 100 times.

To compare the percentage improvement in these scenarios/instances, it is important to ensure that any observed differences are, in fact, due to the instances/scenarios themselves rather than to changes in the "experimental conditions". In our numerical study, the "experimental conditions" are the random numbers required to simulate the demand observed by the retailers. In the following, we outline how the variance reduction technique common random numbers (CRN) was implemented to achieve this. This technique is also called correlated sampling or matched streams. This is discussed further in [10, p. 613].

A single replicate of a particular scenario/instance with horizon length $l$ requires $2l$ random numbers (to simulate the $l$ periods of demand for each of the two retailers). Because each scenario/instance is replicated 100 times, a scenario/instance with horizon length $l$ requires $100 \times 2l$ random numbers. Given that the longest horizon length we consider is $l = 1000$, the longest stream of random numbers we require is 200,000. In those scenarios/instances in which the horizon length is $l = 1000$, this stream of 200,000 random numbers is used. In scenarios/instances in which the horizon length is $l = 10$, for example, the first $100 \times 2 \times 10 = 2000$ random number in this stream are used. That is, in as much as possible, a common stream of random numbers is used for each scenario/instance. To put it differently, for a given horizon length, the same stream of random numbers is used for each combination of the relative size of the retailers' markets, the critical fractile, and the forecast updating method as well as for each

---

1Note that the term precision is not used in the conventional way to describe the reciprocal of the variance but to describe the coefficient of variation of the subjective distribution.
scenario. Moreover, the same stream of random numbers is used to compare the case when the central agent pools the retailers' demand data to the case when the retailers act with local data only.

4.2 Computational Details

4.2.1 Choosing the smoothing parameter and the order of the moving average

To implement the moving average and the exponential smoothing forecast updating methods discussed in §3.5, we must choose a smoothing parameter $\alpha$ and an order for the moving average $N$. In doing this, we seek to make the three updating methods comparable. Besides this, it is desirable that these parameters do not depend on the period $t$. In the following, we justify our approach.

Recall from §3.5.3, the Bayesian forecast after observing demand in period 1 is a weighted average of the mean of the subjective distribution and the demand observed in period 1, where the weights are

$$\frac{(\tau^2)^{-1}}{(\tau^2)^{-1} + (\sigma^2)^{-1}}$$

and

$$\frac{(\sigma^2)^{-1}}{(\tau^2)^{-1} + (\sigma^2)^{-1}}$$

respectively. Also recall that $\tau^2$ is the variance of the subjective distribution and $\sigma^2$ is the variance of demand. Similarly, the exponential smoothing forecast after observing demand in period 1 is also a weighted average of the mean of the subjective distribution and the demand observed in period 1. In this case, however, the weights are $1 - \alpha$ and $\alpha$ respectively. Clearly, these forecasts are equivalent if

$$\alpha \equiv \frac{(\sigma^2)^{-1}}{(\tau^2)^{-1} + (\sigma^2)^{-1}}.$$  \hspace{1cm} (4.1)

Hence, for each scenario, the smoothing parameter $\alpha$ is chosen according to (4.1).

As discussed in [2, pp. 8-10], a small value of $\alpha$ puts relatively more emphasis on old values of demand. The same effect can be achieved when using a moving average forecast by choosing a large value of $N$. In particular, Axsäter [2] proposes the following: The exponential smoothing forecast (after observing the demand in period $t$) is a weighted average of all the previously observed demands, where the weights decrease exponentially with the age of the data. If we assume that we have observed infinite previous demand (i.e., $x_t, x_{t-1}, \ldots$), the "average age" of the data (in periods) is then

$$\alpha \cdot 0 + \alpha(1 - \alpha) \cdot 1 + \alpha(1 - \alpha)^2 \cdot 2 + \ldots = \alpha(1 - \alpha)[1 + 2(1 - \alpha) + 3(1 - \alpha)^2 + \ldots] = \alpha(1 - \alpha)/(1 - (1 - \alpha))^2 = (1 - \alpha)/\alpha.$$

The moving average forecast (after observing the demand in period $t$), is a weighted average of the demands in periods $t$, $t-1$, $t-N+1$. The weights are each $1/N$, and the ages of these data are 0, 1, $\ldots$, $N-1$ periods respectively. The average age of the data is $N(N - 1)/2N = (N - 1)/2$ periods. Thus, a moving average forecast of order $N$ and an exponential smoothing forecast with smoothing parameter $\alpha$ are based on data of the same average age if

$$\frac{(N - 1)/2}{\alpha}.$$

Accordingly, for each scenario, the order of the moving average is chosen to be $N = (2 - \alpha)/\alpha$. Because $N$ must be an integer, we choose $N = \lfloor(2 - \alpha)/\alpha\rfloor$ when implementing this. Table 4.2 gives the values of $\alpha$ and $N$ for each scenario.
4.2.2 Generating uniform random numbers

As mentioned earlier, the numerical study was implemented in C++. Key motivations for this—as opposed to using Microsoft Excel and/or Microsoft Visual Basic for Applications (VBA)—were simulation run-time and random number generation. In this section, we elaborate on why the numerical study was not implemented in Microsoft Excel and/or VBA, and we provide details about the random number generator used.

According to L’Ecuyer [12], a good random number generator runs fast, uses little memory, has a period length which is extremely long, performs the same way regardless of the computing environment, and has the ability to generate the same sequence of numbers. These criteria are not enough, however. Successive calls to the random number generator should be indistinguishable from independent draws from the uniform distribution on the interval (0,1). In a discussion regarding software for uniform random number generation, he stresses the following: “despite repeated warnings over the past years about certain classes of generators, and despite the availability of much better alternatives, simplistic and unsafe [random number generators] still abound in commercial software” [12, §1.1]. In particular, L’Ecuyer discusses two tests to ascertain how closely the numbers produced by a given random number generator resemble true i.i.d. $U(0,1)$ random variates: the birthday spacings test and the collision test. Interestingly, both Microsoft Excel and Microsoft Visual Basic fail these tests even when reasonably few random numbers are drawn. Excel, for instance, fails the birthday spacings test for 16,384 draws, and Visual Basic fails the collision test for 131,072 draws. See [12] for further details.

Warnings such as this were the primary reason for implementing the numerical study in C++. The random number generator used in our numerical study is the maximally equidistributed combined Tausworthe generator proposed by L’Ecuyer [11] and based on a paper by Tausworthe [19]. This random number generator is included in the GNU Scientific Library (GSL) and is recommended for use in simulation because it has an extremely long period, low correlation, and passes most statistical tests [7]. It is related to cryptographic methods and uses a sequence of bits to form random numbers [10, pp. 434-436].

4.2.3 Computing the percentage points of the standard normal distribution

Given a random variate drawn from the uniform distribution on the interval (0,1), the percentage points of the standard normal distribution are needed to simulate demand. That is, they are needed to simulate a random variate from a normal distribution with a given mean and variance. These percentage points are also required to determine the retailers’ newsvendor-optimal order quantities. In our simulation, we use Wichura’s algorithm AS 241 (routine PPND16) [20]. This algorithm computes a numerical approximation to the percentage point $z$ of the standard normal distribution for a given lower tail area $x$ where

$$z = \Phi^{-1}(x),$$
and

\[ \Phi(z) = \int_{-\infty}^{z} (2\pi)^{-1/2} e^{-\frac{y^2}{2}} dy. \]

In the original reference, the algorithm was implemented in Fortran 77. We include a C++ version in Appendix C. According to [20], this routine is accurate to approximately 16 decimals over the range \( \min(x, 1-x) > 10^{-316} \).
Chapter 5

Results

In this chapter we present and interpret the results of our numerical study. First, we compare the value of pooling in the six scenarios we considered. Second, we describe the effect on the value of pooling of the relative size of the retailers’ markets, the critical fractile, the horizon length, and the forecast updating method. Last, we discuss the effect of the horizon length for each forecast updating method. Other interaction effects were investigated, but none were found to be of interest.

5.1 Value of Pooling by Scenario

Recall from §4.1 that each of the 378 instances was replicated 100 times. To summarize the value of pooling in each instance, we found the median percentage improvement in profit with pooling. Because these percentage improvements are ratios, their median was chosen rather than their mean as the measure of central tendency. The percentage improvement in profit with pooling is extremely large—even for small absolute improvements in profit—if the profit without pooling is close to zero. If, for a particular instance, a few such extreme values are observed by chance, the median is less affected than the mean. Table 5.1 gives the five-number summary of the 378 median percentage improvements (%) by scenario.

A number of observations can be made from Table 5.1. In general, there is more value in pooling when the a priori beliefs about mean demand are overconfident (e.g., scenarios III and IV) or underconfident (e.g., scenario V). As expected, there is little value in pooling when the subjective distribution is highly accurate and highly precise (e.g., scenario VI). Regardless of the scenario, there are instances in which there is no value in pooling, and, in fact, there are instances in which the median percentage improvement with pooling is negative. These nearly all occur when the horizon length is very short (e.g., \( l = 2 \)) or very long (e.g., \( l = 500 \) or \( l = 1000 \)). With the exception of scenario V, the median percentage improvement is less than 0.3% in half of the scenarios/instances we considered. The median percentage improvement is less than 2% in 75% of the scenarios/instances we considered (with the exception of scenario V).\(^1\) There are some instances, though, with a considerably higher median percentage improvement with pooling. The maximum value ranges from approximately 6% in scenario II to 150% in scenario IV.

Clearly, very long horizon lengths such as \( l = 500 \) and \( l = 1000 \) are not meaningful in practice; they were included and are valuable because they reveal long-run behaviour. Table 5.2 gives the five-number summary of the median percentage improvements for the 252 instances with horizon length \( l \leq 50 \) by scenario. Surprisingly, there is little change. We will develop further insight into why this is the case when we discuss the effect of the horizon length on the value of pooling.

\(^1\)In scenario V, the median percentage improvement is less than 1% in half of the instances and less than 5% in 75% of the instances.
Chapter 5. Results

Table 5.1: Five-number summary of the median percentage improvements in profit with pooling by scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Min</th>
<th>Q₁</th>
<th>Median</th>
<th>Q₃</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-0.145</td>
<td>0.071</td>
<td>0.278</td>
<td>0.980</td>
<td>7.655</td>
</tr>
<tr>
<td>II</td>
<td>-1.425</td>
<td>0.082</td>
<td>0.280</td>
<td>0.883</td>
<td>6.296</td>
</tr>
<tr>
<td>III</td>
<td>-0.056</td>
<td>0.001</td>
<td>0.172</td>
<td>1.666</td>
<td>20.488</td>
</tr>
<tr>
<td>IV</td>
<td>-0.202</td>
<td>0.028</td>
<td>0.205</td>
<td>1.877</td>
<td>150.488</td>
</tr>
<tr>
<td>V</td>
<td>-0.223</td>
<td>0.223</td>
<td>0.947</td>
<td>4.081</td>
<td>18.439</td>
</tr>
<tr>
<td>VI</td>
<td>-0.076</td>
<td>0.001</td>
<td>0.013</td>
<td>0.111</td>
<td>6.486</td>
</tr>
</tbody>
</table>

Table 5.2: Five-number summary of the median percentage improvements in profit with pooling by scenario for instances with horizon lengths of 50 or less

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Min</th>
<th>Q₁</th>
<th>Median</th>
<th>Q₃</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-0.145</td>
<td>0.100</td>
<td>0.362</td>
<td>1.077</td>
<td>7.655</td>
</tr>
<tr>
<td>II</td>
<td>-1.425</td>
<td>0.113</td>
<td>0.332</td>
<td>0.807</td>
<td>6.296</td>
</tr>
<tr>
<td>III</td>
<td>-0.002</td>
<td>0.001</td>
<td>0.260</td>
<td>1.177</td>
<td>12.270</td>
</tr>
<tr>
<td>IV</td>
<td>-0.202</td>
<td>0.011</td>
<td>0.304</td>
<td>2.781</td>
<td>150.500</td>
</tr>
<tr>
<td>V</td>
<td>-0.223</td>
<td>0.297</td>
<td>0.983</td>
<td>3.876</td>
<td>16.750</td>
</tr>
<tr>
<td>VI</td>
<td>-0.076</td>
<td>0.000</td>
<td>0.009</td>
<td>0.207</td>
<td>6.486</td>
</tr>
</tbody>
</table>

Table 5.3: 95% C.I.'s for the median percentage improvement in profit with pooling by relative size of the retailers' markets and scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Relative Size of the Retailers' Markets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>I</td>
<td>(0.191, 0.215)</td>
</tr>
<tr>
<td>II</td>
<td>(0.201, 0.227)</td>
</tr>
<tr>
<td>III</td>
<td>(0.093, 0.114)</td>
</tr>
<tr>
<td>IV</td>
<td>(0.157, 0.174)</td>
</tr>
<tr>
<td>V</td>
<td>(0.734, 0.857)</td>
</tr>
<tr>
<td>VI</td>
<td>(0.010, 0.012)</td>
</tr>
</tbody>
</table>

5.2 Effect of the Relative Size of the Retailers' Markets

For each value of the relative size of the retailers' markets $k$, we found the median percentage improvement in profit with pooling over all replicates and over all values of the critical fractile, the horizon length, and the forecast updating method. The results (by scenario) are illustrated in Figure 5.1. Table 5.3 gives 95% confidence intervals for the median percentage improvement with pooling by relative size of the retailers' markets and scenario. Clearly, the effect of the relative size of the retailers' markets on the value of pooling is uninteresting. This is likely due to our assumption that when the retailers act with local data only, they share a common subjective distribution: Both retailer 1 and retailer 2 believe $\theta \sim N(\mu, \tau^2)$. We hypothesize that a more interesting effect would be observed if each retailer had his or her own local beliefs about mean demand. This is discussed further in Chapter 6.
Figure 5.1: Effect of the relative size of the retailers' markets on the median percentage improvement in profit with pooling by scenario
### Table 5.4: 95% C.I.’s for the median percentage improvement in profit with pooling by critical fractile and scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Critical Fractile</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.1</td>
<td>0.25</td>
</tr>
<tr>
<td>I</td>
<td>(1.374, 1.522)</td>
<td>(1.001, 1.090)</td>
<td>(0.603, 0.648)</td>
</tr>
<tr>
<td>II</td>
<td>(1.503, 1.738)</td>
<td>(1.277, 1.455)</td>
<td>(0.925, 0.990)</td>
</tr>
<tr>
<td>III</td>
<td>(0.355, 0.437)</td>
<td>(0.244, 0.291)</td>
<td>(0.153, 0.182)</td>
</tr>
<tr>
<td>IV</td>
<td>(1.391, 1.968)</td>
<td>(1.235, 1.806)</td>
<td>(0.843, 1.261)</td>
</tr>
<tr>
<td>V</td>
<td>(7.403, 8.310)</td>
<td>(5.137, 6.045)</td>
<td>(3.343, 3.856)</td>
</tr>
<tr>
<td>VI</td>
<td>(0.072, 0.090)</td>
<td>(0.054, 0.068)</td>
<td>(0.032, 0.039)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Critical Fractile</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.75</td>
<td>0.9</td>
<td>0.95</td>
</tr>
<tr>
<td>I</td>
<td>(0.159, 0.173)</td>
<td>(0.066, 0.073)</td>
<td>(0.029, 0.035)</td>
</tr>
<tr>
<td>II</td>
<td>(0.249, 0.266)</td>
<td>(0.114, 0.121)</td>
<td>(0.060, 0.065)</td>
</tr>
<tr>
<td>III</td>
<td>(0.060, 0.082)</td>
<td>(0.043, 0.060)</td>
<td>(0.032, 0.050)</td>
</tr>
<tr>
<td>IV</td>
<td>(0.084, 0.118)</td>
<td>(0.040, 0.050)</td>
<td>(0.021, 0.026)</td>
</tr>
<tr>
<td>V</td>
<td>(0.840, 0.969)</td>
<td>(0.348, 0.393)</td>
<td>(0.188, 0.206)</td>
</tr>
<tr>
<td>VI</td>
<td>(0.009, 0.011)</td>
<td>(0.005, 0.006)</td>
<td>(0.002, 0.003)</td>
</tr>
</tbody>
</table>

5.3 **Effect of the Critical Fractile**

For each value of the critical fractile \( c \), we found the median percentage improvement in profit with pooling over all replicates and over all values of the relative size of the retailers’ markets, the horizon length, and the forecast updating method. These results are presented (by scenario) in Figure 5.2. Table 5.4 gives 95% confidence intervals for the median percentage improvement with pooling by critical fractile and scenario. As can be seen, the percentage improvement increases as the critical fractile decreases. For reasonable critical fractiles (i.e., \( c = 0.5 \), \( c = 0.75 \), and \( c = 0.9 \)), there is little value in pooling (except for in scenario V). For unrealistically low critical fractiles (i.e., \( c = 0.05 \) and \( c = 0.1 \)), the value of pooling is moderate in scenarios I, II and IV, and it is considerable in scenario V.

Recall from §4.1 that for a critical fractile \( c \), we chose the selling price \( p = 1 + c \), the wholesale cost \( w = 1 \) and the salvage value \( s = c \). Given this choice, the underage cost (marginal profit) is always equal to the salvage value—regardless of the value of \( c \). Specifically, \( p - w = s = c \). Hence, calling the effect of the parameter \( c \) the “effect of the critical fractile” is somewhat misleading. It is the effect of the critical fractile but only in this particular setting. This is addressed in greater detail in Chapter 6.

In light of our choice of the cost parameters, however, the observed increase in the value of pooling as the critical fractile decreases can be explained as follows: As the value of \( c \) decreases, the underage cost \( p - w \) decreases. The firm’s potential profit in a period (given the observed demands) also decreases. That is, if in period \( t \) retailer \( i \) orders quantity \( q_{t,i} \), the firm’s maximum profit,

\[
\sum_{t=1,2} (p - w)d_{t,i},
\]

is realized if \( d_{t,i} = q_{t,i} \) for \( i = 1, 2 \). Consequently, as \( p - w \) decreases, the firm’s potential profit decreases. When the underage cost is low, accurately forecasting demand is crucial to realizing as much of this potential profit as possible. Hence, the value of information is greater.
Figure 5.2: Effect of the critical fractile on the median percentage improvement in profit with pooling by scenario
5.4 Effect of the Horizon Length

For each horizon length \( t \), we found the median percentage improvement in profit with pooling over all replicates and over all values of the relative size of the retailers' markets, the critical fractile, and the forecast updating method. The results (by scenario) are shown in Figure 5.3. Table 5.5 gives 95% confidence intervals for the median percentage improvement with pooling by horizon length and scenario. Generally, the value of pooling is greater for moderate horizon lengths than for short or for long horizon lengths.

The observed decrease in the value of pooling as the horizon length increases is consistent with intuition and can be explained as follows: When the horizon length is long, the retailers acquire more information about the demand distribution (acting with local data) than they do for short horizon lengths. The additional information the central agent gains by pooling the retailers' demand data is less beneficial because it provides little new information. If the central agent learns the demand distribution, pooling the retailers' data provides no new information; consequently, there is no longer any benefit to pooling. Recall that we measure the percentage improvement in the average profit over the horizon when the central agent pools the retailers' demand data versus the average profit over the horizon when the retailers act with local data only. When the horizon length is long (and there are many periods in which there is little or no benefit to pooling), the average profit with pooling and the average profit without pooling are nearly equal. Thus, the percentage improvement is negligible.

Based on this analysis, it seems counter-intuitive that the value of pooling does not strictly decrease as the horizon length increases. We propose two reasons for why this was not seen. First, the firm's profits are the same in the initial period regardless of whether the central agent pools the retailers' demand data or the retailers' act with local data only: When constructing the predictive distribution in the first period, no demand has occurred; thus, the same ordering decisions are made, and, since the same demand is observed, the same profit is realized. This "initial period effect" decreases the apparent value of pooling for short horizon lengths more than for long horizon lengths. Second, and most important, there is a notable interaction between the effect of the horizon length and the effect of the forecast updating method on the value of pooling. This will be discussed in §5.6. In light of this, interpreting the main effect of the horizon length is difficult.

5.5 Effect of the Forecast Updating Method

For each forecast updating method, we found the median percentage improvement in profit with pooling over all replicates and over all values of the relative size of the retailers' markets, the critical fractile, and the horizon length. These findings are summarized (by scenario) in Figure 5.4. Table 5.6 gives 95% confidence intervals for the median percentage improvement with pooling by forecast updating method and scenario. The effect of the forecast updating method on the value of pooling is similar in scenarios I and II as well as in scenarios III and IV. This is expected since in scenarios I and II, the subjective distribution is reasonably accurate and reasonably precise. In scenarios III and IV, on the other hand, the subjective distribution is very inaccurate but highly precise—the \( \text{a priori} \) beliefs about mean demand are overconfident.

When the \( \text{a priori} \) beliefs about mean demand are reasonable (e.g., scenarios I and II), pooling is most valuable when a moving average forecast update is used; pooling is least valuable when a Bayesian forecast update is used. Regardless of the forecast updating method, however, the median percentage improvement is marginal. When the \( \text{a priori} \) beliefs about mean demand are underconfident (e.g., scenario V), pooling is also most valuable when a moving average forecast update is used and least valuable when a Bayesian forecast update is used. In particular, the median percentage improvement is moderate in both the moving average and the exponential smoothing cases but negligible in the Bayesian case. When the \( \text{a priori} \) beliefs about mean
Figure 5.3: Effect of the horizon length on the median percentage improvement in profit with pooling by scenario
Table 5.5: 95% C.I.’s for the median percentage improvement in profit with pooling by horizon length and scenario

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Horizon Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>I</td>
<td>(0.121, 0.165)</td>
</tr>
<tr>
<td>II</td>
<td>(0.096, 0.128)</td>
</tr>
<tr>
<td>III</td>
<td>(0.067, 0.080)</td>
</tr>
<tr>
<td>IV</td>
<td>(0.097, 0.148)</td>
</tr>
<tr>
<td>V</td>
<td>(0.291, 0.360)</td>
</tr>
<tr>
<td>VI</td>
<td>(0.001, 0.002)</td>
</tr>
</tbody>
</table>

Table 5.5: 95% C.I.’s for the median percentage improvement in profit with pooling by horizon length and scenario

demand are overconfident (e.g., scenarios III and IV), pooling is most valuable when a Bayesian forecast update is used; pooling is least valuable when an exponential smoothing forecast update is used. Furthermore, the median percentage improvement is moderate in the Bayesian case but negligible in both the moving average and exponential smoothing cases. Finally, as expected, the median percentage improvement is negligible when the subjective distribution is highly accurate and highly precise (e.g., scenario VI).

Recall from §4.2.1 that the parameters of the subjective distribution determine the order of the moving average $N$, the smoothing parameter $\alpha$, and the prior distribution in the moving average, exponential smoothing, and Bayesian forecast updating methods respectively. Given this relationship, we now discuss how the scenario affects the value of pooling for each forecast updating method.

Moving average. When a moving average forecast update is used, the value of pooling depends on the order of the moving average. From Table 4.2, we chose the following values in scenarios I through VI respectively: $N = 12$, $N = 7$, $N = 800$, $N = 89$, $N = 1$, and $N = 200$. From Figure 5.4 and Table 5.6, the smaller the value of $N$, the more value there is in pooling. The median percentage improvement in scenario III (with $N = 800$), for example, is 0.134%; in scenario V (with $N = 1$) it is 2.022%. This result is intuitive: Given that demand is stationary, the best estimate of mean demand is the average of all the observed demands. In the extreme case when the order of the moving average $N$ is longer than the horizon length $I$, the moving average forecast is indeed an average of all the previous demands. In this setting—for large $I$—there is little benefit to pooling the retailers’ demand data since, even without pooling, the demand distribution is essentially known.

Exponential smoothing. When an exponential smoothing forecast update is used, the value of pooling depends on the smoothing parameter. From Table 4.2, we chose the following values in scenarios I through VI respectively: $\alpha = 0.153$, $\alpha = 0.248$, $\alpha = 0.002$, $\alpha = 0.022$, $\alpha = 0.800$, and $\alpha = 0.010$. From Figure 5.4 and Table 5.6, the larger the smoothing parameter $\alpha$, the more
value there is in pooling. In scenarios III, IV and VI (with \( \alpha = 0.002, \alpha = 0.022, \) and \( \alpha = 0.010 \) respectively) the benefit of pooling is negligible; the median percentage improvement in each scenario is less than 0.006%. In scenario V (with \( \alpha = 0.800 \)) the median percentage improvement is 1.409%. We suggest the following interpretation: From the non-recursive equation given in (B.1), the exponential smoothing forecast is a weighted average of all previously observed demands, where the weights decrease exponentially with the age of the data. When the smoothing parameter \( \alpha \) is small, relatively more emphasis is put on older values of demand than when \( \alpha \) is large. Small values of \( \alpha \) also put relatively more emphasis on the initial forecast (i.e., the mean of the subjective distribution \( \mu \)); thus, there is little benefit to pooling because the forecast is largely determined by \( \mu \). Conversely, for large values of \( \alpha \), the forecast is determined to a greater extent by the observed demands. Hence, the additional information the central agent gains by pooling the retailers' demand data is more valuable.

**Bayesian updating.** When a Bayesian forecast update is used, the value of pooling depends on the parameters of the prior distribution \( \mu \) and \( \tau^2 \). From Figure 5.4 and Table 5.6, the value of pooling is the greatest when the \( \text{a priori} \) beliefs about mean demand are overconfident (e.g., scenarios III and IV). The median percentage improvements in scenarios III and IV are 2.118% and 2.302% respectively. The percentage improvement is negligible in all other scenarios. In the following, we illustrate why this is the case.

Recall from §3.5.3 that the Bayesian forecast update (after observing demand in period \( t \)) is a weighted average of the demand observed in period \( t \) and the mean of the posterior distribution in period \( t - 1 \). For convenience, we define \( w_{t,p} \) and \( w_t \) to be the ratio of these weights when the central agent pools the retailers' demand data and when the retailers act with local data respectively. That is, \( w_{t,p} = 2\tau^2_{t-1,p}/\tau^2 \) and \( w_t = \tau^2_{t-1}/\sigma^2 \), where \( \tau^2_{t-1,p} \) and \( \tau^2_{t-1} \) are the pooled and non-pooled variances of the posterior distributions in period \( t - 1 \), and \( \sigma^2 \) is the variance of demand. Also, we define

\[
W_t \equiv \frac{w_{t,p}}{w_t}.
\]

In the first period, \( W_1 = 2\tau^2/\tau^2 = 2 \); as \( t \) approaches infinity, \( W_t \) approaches unity.\(^2\) Moreover, for small values of \( \tau^2 \), the convergence is slower. To illustrate, it can be verified numerically that in scenario II (with \( \tau^2 = 5.29 \)) we have \( W_1 = 2, W_2 = 1.602, W_5 = 1.274, W_{10} = 1.144, \) and \( W_{50} = 1.030 \). In scenario III (with \( \tau^2 = 0.04 \)) we have \( W_1 = 2, W_2 = 1.995, W_5 = 1.990, W_{10} = 1.957, \) and \( W_{50} = 1.803 \). In summary, for small values of \( \tau^2 \), more weight is placed on the mean of the posterior distribution than on the data. When the central agent pools the retailers' demand data, relatively more weight is given to the data than when the retailers act with local data. Specifically, the data receive twice as much weight in the pooled than non-pooled case in period 1, but this decreases to unity as \( t \) approaches infinity. Because the convergence is slower for small values of \( \tau^2 \), more can be learned about the demand distribution when the central agent pools the retailers' demand data if \( \tau^2 \) is small. Hence, the value of pooling is greater.

It should be noted, however, that the value of pooling also depends on the accuracy of the subjective distribution. When \( \tau^2 \) is small, but the subjective distribution is highly accurate, there is little value in pooling as no learning is needed. The median percentage improvement in scenario IV (with \( \tau^2 = 0.36 \)) is 2.302%. In contrast, in Scenario VI (with \( \tau^2 = 0.16 \), the median percentage improvement is only 0.002% because \( \mu/\theta = 1 \).

\(^2\)This was investigated numerically. As \( t \) approaches infinity, \( \tau^2_{t-1,p} \) and \( \tau^2_{t-1} \) approach zero, but \( \tau^2_{t-1,p} \) converges faster.
Figure 5.4: Effect of the forecast updating method on the median percentage improvement in profit with pooling by scenario

Table 5.6: 95% C.I.’s for the median percentage improvement in profit with pooling by forecast updating method and scenario
5.6 Effect of the Horizon Length by Forecast Updating Method

For each forecast updating method and horizon length $l$, we found the median percentage improvement in profit with pooling over all replicates and over all values of the relative size of the retailers' markets and the critical fractile. These results are presented in Figures 5.5 through 5.7. Clearly, the effect of the horizon length on the value of pooling depends on the forecast updating method.

Moving average. In §5.5 we illustrated that when a moving average forecast update is used, the value of pooling depends on the order of the moving average $N$. We will now show that the effect of the horizon length also depends on $N$. See Figure 5.5. Notably, in scenarios III, IV and VI we see the horizon length effect discussed in §5.4: The value of pooling is greater when $l = 5$ than when $l = 2$ (because of the effect of the initial period), but as the horizon length becomes longer it decreases. In contrast, in scenario V, the value of pooling increases as the horizon length increases.

Scenario V is a particularly special case when a moving average forecast update is used. Recall from Table 4.2 that in this scenario the order of the moving average is $N = 1$. We
Figure 5.6: Effect of the horizon length on the median percentage improvement in profit with pooling by scenario when an exponential smoothing forecast update is used.
Figure 5.7: Effect of the horizon length on the median percentage improvement in profit with pooling by scenario when a Bayesian forecast update is used.
anticipated the horizon length would not affect the value of pooling in this case. Our reasoning is as follows: We do not expect the value of pooling to decrease as the horizon length increases because the demand distribution is never learned. To put it differently, the parameters of the predictive distribution never converge to the parameters of the demand distribution: The forecast reacts to the demand observed in the previous period, and the variance of the forecast is constant each period (see §3.5.1). Also, we do not expect the value of pooling to increase as the horizon length increases because, regardless of the period, the reduction in the variance of the forecast with pooling is constant. As will be seen later, this is not the case when an exponential smoothing forecast update is used. Again, we propose the effect of the initial period as an explanation for why the observed percentage improvement with pooling increases as the horizon length increases. It should be noted that this initial period effect is present in all scenarios though.

It is also interesting that in scenarios III, IV and VI, the percentage improvement is nearly 0% when $l = 500$. In scenarios I and II, on the other hand, the percentage improvement is marginal even when $l = 1000$. This is consistent with the order of the moving average in each scenario. Recall that in scenarios III, IV and VI, the values of $N$ are 800, 89, and 200 respectively. For such large values of $N$—when the horizon length is long—the information the data can provide is exhausted even without pooling; thus, there is no additional benefit to pooling. In scenarios I and II, the values of $N$ are 12 and 7 respectively, so it is reasonable that pooling is beneficial even for very long horizon lengths.

**Exponential smoothing.** Also in §5.5, we showed that when an exponential smoothing forecast update is used, the value of pooling depends on the smoothing parameter. The percentage improvement with pooling is negligible for extremely small values of $\alpha$ (e.g., scenarios III, IV and VI with $\alpha = 0.002$, $\alpha = 0.022$ and $\alpha = 0.010$ respectively), marginal for reasonable values of $\alpha$ (e.g., scenarios I and II with $\alpha = 0.153$ and $\alpha = 0.248$ respectively), and modest for large values of $\alpha$ (e.g., scenario V with $\alpha = 0.800$). In scenarios where there is value in pooling, the magnitude is affected by the horizon length. This can be seen in Figure 5.6. Generally, the value of pooling increases as the horizon length increases. The effect is greater in scenarios where the value of pooling is greater. In the following, we explain this result.

Recall that the variance of the pooled exponential smoothing forecast (after observing demand in period $t$) is given by

$$V_{t,p} = \frac{\alpha}{2 - \alpha} \left(1 - \left(1 - \alpha\right)^{2t}\right) \frac{\sigma^2}{2} + \left(1 - \alpha\right)^{2t} \sigma^2,$$

and the variance of the non-pooled forecast is given by

$$V_t = \frac{\alpha}{2 - \alpha} \left(1 - \left(1 - \alpha\right)^{2t}\right) \sigma^2 + \left(1 - \alpha\right)^{2t} \sigma^2.$$

The relative efficiency of the pooled versus the non-pooled forecast is defined to be the ratio of the variance of the pooled forecast to the variance of the non-pooled forecast. As $t$ approaches infinity, the variance of the pooled forecast decreases, approaching

$$\left(\frac{\alpha}{2 - \alpha}\right) \frac{\sigma^2}{2}.$$

Similarly, the variance of the non-pooled forecast decreases, approaching

$$\left(\frac{\alpha}{2 - \alpha}\right) \sigma^2.$$

Hence, as $t$ approaches infinity, the relative efficiency decreases, approaching $1/2$. Pooling is more valuable the closer the relative efficiency is to $1/2$. In other words, there is more value
Table 5.7: Relative efficiency of the pooled versus the non-pooled exponential smoothing forecast in a sample of periods for scenarios I, II, and V

<table>
<thead>
<tr>
<th>period</th>
<th>Scenario</th>
<th>I</th>
<th>II</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.923</td>
<td>0.876</td>
<td>0.600</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.849</td>
<td>0.762</td>
<td>0.505</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.669</td>
<td>0.563</td>
<td>0.500</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.538</td>
<td>0.504</td>
<td>0.500</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>0.500</td>
<td>0.500</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Table 5.7 gives the relative efficiency in a sample of periods for scenarios I, II, and V. This table also justifies the observed differences in the percentage improvement with pooling between these scenarios. In scenarios I and II, the percentage improvement is marginal but approximately 30 times greater when \( t = 1000 \) than when \( t = 2 \). In scenario V, the percentage improvement is moderate and less than 10 times greater when \( t = 1000 \) than when \( t = 2 \). This is because in scenarios I and II, the relative efficiency in the first period is close to 1, implying there is initially little benefit to pooling, so much can be gained over time. In scenario V, the initial relative efficiency is 0.6, implying there is already considerable benefit to pooling and less to be gained over time.

Bayesian updating. As we have already seen, when a Bayesian forecast update is used, the value of pooling depends on the parameters of the prior distribution \( \mu \) and \( \tau^2 \). There is considerable value in pooling when the \textit{a priori} beliefs about mean demand are overconfident but little value in pooling otherwise. In addition, Figure 5.7 shows that for each scenario there is a horizon length \( I^* \) for which, when the horizon length is shorter, the value of pooling increases as the horizon length increases and for which, when the horizon length is longer, the value of pooling decreases as the horizon length increases.

The observed decrease in the value of pooling as the horizon length increases (for horizon lengths greater than \( I^* \)) can be explained as follows: When a Bayesian forecast update is used—unlike in the case of a moving average or an exponential smoothing forecast update—the parameters of the predictive distribution do, in fact, converge to the parameters of the demand distribution. After a number of periods (if the horizon length is long enough), the demand distribution will be known, and there will no longer be any value in pooling the retailers' demand data. Consequently, the value of pooling decreases as the horizon length increases. Small values of \( \tau^2 \) relative to \( \sigma^2 \) put more weight on the mean of the posterior distribution than on the observed demand in the forecast update (see §3.5.3); therefore, this convergence is slower (i.e., the value of \( I^* \) is greater). To illustrate, \( I^* \) is greater in scenarios III and IV (with \( \tau^2 = 0.04 \) and \( \tau^2 = 0.36 \) respectively) than in scenarios I, II, and V (with \( \tau^2 = 2.89, \tau^2 = 5.29, \) and \( \tau^2 = 64.00 \) respectively). In scenario VI, the \textit{a priori} beliefs about mean demand are highly accurate and highly precise; the demand distribution does not need to be learned, so there is no value in pooling regardless of the horizon length. The observed increase in the value of pooling as the horizon length increases (for horizon lengths shorter than \( I^* \)), however, is unexpected. Once again, we hypothesize that this is due to the effect of the initial period.
Chapter 6

Conclusions

In this paper we have reported on a numerical study that investigates the value in pooling demand data across a single firm's multiple retail locations. In particular, we considered the effect of the relative size of the retailers' markets, the critical fractile, the horizon length, and the demand forecast updating method. Furthermore, we studied a variety of scenarios characterizing the \textit{a priori} beliefs about mean demand. Our successes are numerous and worth highlighting. First, we chose a meaningful objective—there is a growing interest in the area of information sharing, particularly information sharing between a retailer and an upstream supplier. There is little (if any) research, however, which considers information sharing between a firm's retailers. Second, we constructed a valid model: The newsvendor problem is often studied in the supply chain literature; we thoughtfully decided how to initialize each forecast updating method; and, for each scenario, we carefully chose the value of the smoothing parameter and the order of the moving average. Last, the numerical experiment was well-designed. We selected a meaningful range of levels for each factor, as well as certain extreme levels to learn about the value of pooling under these atypical conditions. Also, we used a simulation-quality random number generator and applied a common random numbers variance reduction technique.

Many of the results we obtained were expected. For instance, there is more value in pooling the retailers' demand data when the \textit{a priori} beliefs about mean demand are overconfident or underconfident. From a managerial perspective, a significant result is that regardless of the \textit{a priori} beliefs, pooling is more beneficial when a moving average forecast update is used than when an exponential smoothing forecast update is used. This is of interest because a moving average and exponential smoothing are forecasting techniques commonly used in practice.

It is surprising though that, on the whole, we found little value in pooling in the majority of the scenarios/instances we studied. Recall from Table 5.2 that the percentage improvement is less than 0.4% in 50% of the instances with meaningful horizon lengths (in all but scenario V). The percentage improvement is less than 4% in 75% of the instances. Furthermore, it is noteworthy that if we restrict our attention to those instances in which the critical fractile is 0.75 or greater and the horizon length is 50 or less, pooling is only beneficial when the \textit{a priori} beliefs about mean demand are overconfident or underconfident. In particular, in scenarios I through VI respectively, the largest median percentage improvement with pooling is 0.60%, 0.60%, 3.42%, 2.62%, 1.72%, and 0.59%. In each of scenarios III and IV, this maximum occurs when a Bayesian forecast update is used. In other words, even for reasonable critical fractiles and meaningful horizon lengths, there is value in pooling in the Bayesian case when the \textit{a priori} beliefs are overconfident.

In the following, we summarize the effects of the factors we investigated. The relative size of the retailers' markets does not appear to have an interesting effect on the value of pooling. The value of pooling is greater when the critical fractile is smaller. The effect of the horizon length on the value of pooling depends on the forecast updating method. Notwithstanding, we discussed a number of explanations for a change in percentage improvement with a change in horizon length: the effect of the initial period; whether or not the demand distribution is learned over time; the weight given to the observed demand in the pooled versus the non-pooled forecast; and, the relative efficiency of the pooled versus the non-pooled forecast. Finally, the

\footnote{In scenario V, the percentage improvement is less than 1% in 50% of the instances.}
effect of the forecast updating method depends on the scenario (i.e., subjective distribution) as this determines the order of the moving average, the smoothing parameter and the parameters of the prior distribution.

Throughout this paper, we highlighted a number of modelling decisions which warrant further consideration. One such decision is how to choose the cost parameters to achieve a desired critical fractile. The critical fractile does not uniquely determine the selling price $p$, the wholesale cost $w$, or the salvage value $s$. Namely, for a given value $c$, the critical fractile $c = (p - w)/(p - s)$ specifies a single equation with three unknown parameters; we must make two assumptions about the values of these parameters before all three can be uniquely determined. There are many reasonable ways of doing this, and we chose $p = 1 + c$, $w = 1$, and $s = c$. In the following, we outline our reasoning: If we hold $p - s$ constant and we increase (decrease) $p - w$, the critical fractile increases (decreases). When $p - s = 1$, we must have $p - w = c$. That is, $p = s + 1$ and $p = c + w$ which implies $s + 1 = c + w$. We still have one equation with two unknown parameters, so there is not a unique solution. A simple choice, however, is $s = c$ and $w = 1$. In other words, $p = 1 + c$, $w = 1$, and $s = c$. Given this choice, the critical fractile is $c$. As mentioned in §5.3, regardless of the value of $c$, the underage cost is always equal to the salvage value (with $p - w = s = c$). Clearly, there are many alternative approaches to choosing these parameters. One alternative is to study the effects of the underage cost and the overage cost on the value of pooling separately. Fixing the value of $w$—and specifying the underage cost and the overage cost—would uniquely determine $p$, $w$, and $s$, and the critical fractile could be computed from these.

A second issue we must address is our assumption that when the retailers act using local data only they share a common subjective distribution. It would be more meaningful, and probably more interesting, if the retailers had local subjective distributions.

Although we have discovered interesting results, we undertook this numerical study knowing that it would be primarily valuable as a foundation for further research. Many aspects of our model suggest direction for this future work. First, we assumed that the demand observed by the retailers is stationary. How would non-stationary demand affect our results? Specifically, it would be of interest to study how non-stationary demand would impact the effect of the forecast updating method on the value of pooling. Second, we chose the normal distribution for demand and the a priori beliefs about mean demand. In this case, closed-form Bayesian updating formulas are known and easily implemented. This choice, however, is unnatural; it leads to a positive probability of observing negative demand and suggests that there is a positive probability that mean demand is negative. Furthermore, because the predictive distribution is also normal, it allows negative newsvendor-optimal order quantities. Alternatively, we plan to consider an exponential demand distribution with a gamma distribution describing the a priori beliefs about mean demand. This is described in [9]. Third, we assumed that the retailers observe actual demand, that is, their demand is not censored by their order quantities. In future research, we plan to study the censored demand case: When a retailer's order quantity in a period exceeds demand, actual demand is observed; if fewer units are ordered than are demanded, only the number of units sold is observed. Clearly, the Bayesian updating formulas are complicated by this assumption. This is also discussed in [9]. We anticipate that the value of pooling is considerably greater in the censored demand case than in the uncensored case.

2 Recall that no negative demands were actually observed.

3 No order was placed in these cases.
Bibliography


Appendix A

Sampling Distribution of the Moving Average Forecast

Recall from §3.5.1, when the retailers act with local data only, their local exponential smoothing forecasts at the end of period $t$, for $t \geq N$, are

$$\hat{\theta}_{t,i} = \frac{1}{N} \sum_{j=1}^{N} x_{t-j+1,i}, \text{ for } i=1,2.$$  

From §3.1, $X_{t,1}$ and $X_{t,2}$, for $t = 1, 2, \ldots, l$, are i.i.d. $N(\theta, \sigma^2)$. Thus, the mean of the sampling distribution of $\hat{\theta}_{t,i}$, for $t \geq N$, is

$$E[\hat{\theta}_{t,i}] = \frac{1}{N} \sum_{j=1}^{N} N \text{E}[X_{t-j+1,i}]$$

$$= \frac{N \theta}{N}$$

$$= \theta,$$

and the variance is

$$\text{Var}[\hat{\theta}_{t,i}] = \frac{1}{N^2} \sum_{j=1}^{N} N \text{Var}[X_{t-j+1,i}]$$

$$= \frac{N^2 \sigma^2}{N^2}$$

$$= \frac{\sigma^2}{N}. \quad (A.1)$$

Note that when the central agent pools the retailers' demand data, $\sigma^2/2$ replaces $\sigma^2$ in (A.1).
Appendix B

Sampling Distribution of the Exponential Smoothing Forecast

Recall from §3.5.2, when the retailers act with local data only, their local exponential smoothing forecasts at the end of period \( t \), for \( t \geq 1 \), are

\[
\hat{\theta}_{t,i} = \alpha x_{t,i} + (1 - \alpha)\hat{\theta}_{t-1,i}, \quad \text{for } i = 1, 2,
\]

with \( \hat{\theta}_{0,i} = \mu \) for \( i = 1, 2 \). Recursively substituting for the previous forecasts, we obtain the following:

\[
\hat{\theta}_{t,i} = \alpha \sum_{j=0}^{t-1} (1 - \alpha)^j x_{t-j,i} + (1 - \alpha)^t \mu, \quad \text{for } i = 1, 2 \text{ and } t \geq 1. \tag{B.1}
\]

From §3.1, \( X_{t,1} \) and \( X_{t,2} \), for \( t = 1, 2, \ldots, l \), are i.i.d. \( N(\theta, \sigma^2) \). Thus, the mean of the sampling distribution of \( \hat{\theta}_{t,i} \), for \( t \geq 1 \), is

\[
E[\hat{\theta}_{t,i}] = \alpha \sum_{j=0}^{t-1} (1 - \alpha)^j E[X_{t-j,i}] + (1 - \alpha)^t \mu
\]

\[
= \alpha \theta \sum_{j=0}^{t-1} (1 - \alpha)^j + (1 - \alpha)^t \mu
\]

\[
= \alpha \left( \frac{1 - (1 - \alpha)^t}{1 - (1 - \alpha)} \right) \theta + (1 - \alpha)^t \mu, \tag{B.2}
\]

and the variance is

\[
\text{Var}[\hat{\theta}_{t,i}] = \alpha^2 \sum_{j=0}^{t-1} (1 - \alpha)^j \text{Var}[X_{t-j,i}] + (1 - \alpha)^{2t} \sigma^2
\]

\[
= \alpha^2 \sigma^2 \sum_{j=0}^{t-1} (1 - \alpha)^j + (1 - \alpha)^{2t} \sigma^2
\]

\[
= \frac{\alpha}{2 - \alpha} (1 - (1 - \alpha)^{2t}) \sigma^2 + (1 - \alpha)^{2t} \sigma^2. \tag{B.3}
\]

Brown [3, pp. 101-110] defines the exponential smoothing forecast after observing demand in period \( t \) as a linear combination of infinite past observations (i.e., \( x_t, x_{t-1}, \ldots \)), and he shows that \( E[\hat{\theta}_{t,i}] = \theta \) and \( \text{Var}[\hat{\theta}_{t,i}] = \frac{\alpha}{2 - \alpha} \sigma^2 \). These are consistent with (B.2) and (B.3) as \( t \) approaches infinity. Note that when the central agent pools the retailers’ demand data, \( \sigma^2/2 \) replaces \( \sigma^2 \) in (B.3).
Appendix C

Algorithm AS 241

In the following we give a C++ version of Wichura’s algorithm AS 241 (routine PPND16) [20].

```cpp
double ppnd16(double prob) {
    double q, r, ppnd16;
    const double Zero = 0.0E+00;
    const double One = 1.0E+00;
    const double Half = One/2.0E+00;
    const double Split1 = 0.425E+00;
    const double Split2 = 5.0E+00;
    const double Const1 = 0.180625E+00;
    const double Const2 = 1.6E+00;
    // Coefficients for p close to 0.5
    const double a0 = 3.3871328727963666080E+00;
    const double a1 = 1.3314166789178437745E+02;
    const double a2 = 1.9715909503065514427E+03;
    const double a3 = 1.373169376550946125E+04;
    const double a4 = 4.5921953931549871457E+04;
    const double a5 = 6.7265770927008700853E+04;
    const double a6 = 3.3430575583588128105E+04;
    const double a7 = 2.5090802928730122672E+03;
    const double b1 = 4.23133070160911252E+01;
    const double b2 = 6.8718700749205790830E+02;
    const double b3 = 5.394160214247511077E+03;
    const double b4 = 2.1213794301586595867E+04;
    const double b5 = 3.907895800092710610E+04;
    const double b6 = 2.8729085735721942674E+04;
    const double b7 = 5.2264952788528545610E+03;
    // Coefficients for p not close to 0, 0.5 or 1
    const double c0 = 1.42343711074968357734E+00;
    const double c1 = 4.63033784165564529590E+00;
    const double c2 = 5.76949722146069140550E+00;
    const double c3 = 3.64784832476320460504E+00;
    const double c4 = 1.2704582524523683258E+00;
    const double c5 = 2.41780725177450611770E-01;
    const double c6 = 2.27238449892691845833E-02;
    const double c7 = 7.74545014278341407640E-04;
    const double d1 = 2.0531916263775882187E+00;
    const double d2 = 1.67638483018380384940E+00;
    return ppnd16;
}
```
const double d3 = 6.89767334985100004550E-01;
const double d4 = 1.48103976427480074590E-01;
const double d5 = 1.5198665663164571966E-02;
const double d6 = 5.47593808499534494600E-04;
const double d7 = 1.05075007164441684324E-09;

// Coefficients for p near 0 or 1
const double eO = 6.665790464350110377720E+00;
const double e1 = 5.46378491116411436990E+00;
const double e2 = 1.78482653991729133580E+00;
const double e3 = 2.9650571828504891230E-01;
const double e4 = 2.6532189526561230930E-02;
const double e5 = 1.24266094738807843860E-03;
const double e6 = 2.71155556874348757815E-05;
const double e7 = 2.01033439929228813265E-07;
const double f1 = 5.99832206555887937690E-01;
const double f2 = 1.36929880922735805310E-01;
const double f3 = 1.48753612908506148525E-02;
const double f4 = 7.86869131145613259100E-04;
const double f5 = 1.84631831751005468180E-05;
const double f6 = 1.42151175831644588870E-07;
const double f7 = 2.0442631033899378564E-15;

q = prob - Half;

if (fabs(q) <= Split1) {
    r = Const1 - q*q;
    ppndl6 = q*(((a7*r + a6)*r + a5)*r + a4)*r + a3)*r + a2)*r
             + a1)*r + a0)/((((b7*r + b6)*r + b5)*r + b4)*r + b3)*r + b2)*r
             + b1)*r + One);
    return ppndl6; }

else {
    if (q < 0)
        r = prob;
    else
        r = One - prob;

    if (r <= 0) {
        ppndl6 = Zero;
        return ppndl6; }

    r = sqrt(-log(r));

    if (r <= Split2) {
        r = r - Const2;
        ppndl6 = (((((c7*r + c6)*r + c5)*r + c4)*r + c3)*r + c2)*r
                   + c1)*r + c0)/((((d7*r + d6)*r + d5)*r + d4)*r + d3)*r + d2)*r
                   + d1)*r + One); }

else
else {
    r = r - Split2;
    ppnd16 =((((e7*r + e6)*r + e5)*r + e4)*r + e3)*r + e2)*r
    + e1)*r + e0)/((((f7*r + f6)*r + f5)*r + f4)*r + f3)*r + f2)*r
    + f1)*r + One); }

if (q<0)
    ppnd16 = - ppnd16;

return ppnd16; }
Appendix D

Confidence Interval for a Median

For i.i.d. observations $X_1, X_2, \ldots, X_n$ from a continuous distribution, a confidence interval for the population median $\eta$ is of the form

$$(X(k), X(n-k+1)),$$  \hspace{1cm} (D.1)

where $X(k)$ is the $k$th order statistic. The coverage probability of this interval is

$$P(X(k) \leq \eta \leq X(n-k+1)) = 1 - P(\eta < X(k)) - P(\eta < X(n-k+1)).$$

If $\eta < X(k)$, then at most $k - 1$ observations are less than $\eta$. Similarly, if $\eta < X(n-k+1)$, then at most $k - 1$ observations are greater than $\eta$. Thus,

$$P(\eta < X(k)) = \sum_{j=0}^{k-1} P(j \text{ observations are less than } \eta),$$

and

$$P(\eta < X(n-k+1)) = \sum_{j=0}^{k-1} P(j \text{ observations are greater than } \eta).$$

By definition, the probability that an observation is less than $\eta$ or greater than $\eta$ is $\frac{1}{2}$. Thus, the coverage probability is

$$1 - \frac{1}{2^{n-1}} \sum_{j=0}^{k-1} \binom{n}{j}.$$ 

The confidence interval in (D.1) is exact, and it relies only on the assumption that the distribution is continuous and the observations are independent. Further details can be found in [18].

The R/S-PLUS function `median.CI(x, coverage=0.95)` computes a confidence interval for a population median given a numeric vector $x$ and a coverage probability $coverage$. If the length of the vector $x$ is non-zero, a vector containing the lower and upper limits of the confidence interval is returned. If the length of the vector $x$ is zero, 'NA' is returned for each. The default coverage is 0.95, that is, a 95% confidence interval for the population median is computed. The source code is given below:

```r
median.CI <- function(x, coverage=0.95) {
  n <- length(x)
  epsilon <- 1e-12
  if (n > 0) {
    prob <- 0.5
    p <- (1-coverage)/2
    if (abs(p-pbinom(qbinom(p,n,prob),n,prob)) < epsilon) {
      k <- qbinom(p,n,prob)+1
    } else {
    }
```
Appendix D. Confidence Interval for a Median

```r
k <- qbinom(p,n,prob)
if (k>0) {
  lower.limit <- sort(x)[k]
  upper.limit <- sort(x)[n-k+1]
} else {
  lower.limit <- NA
  upper.limit <- NA
}
else {
  lower.limit <- NA
  upper.limit <- NA
}
confidence.interval <- as.data.frame(list(lower=lower.limit,
                                         upper=upper.limit))
confidence.interval
```