Chiao Da Management Review
Vol. 34 No.1, 2014
pp.177-199

報童問題中以不等加權方式組合需求預測值

Combining Demand Forecasts in a Newsboy Problem Using an Unequally Weighted Method

李智明[1] Chih-Ming Lee
東吳大學 企業管理學系
Department of Business Administration, Soochow University
萬文隆 Wen-Lung Wan
東吳大學 企業管理學系
Department of Business Administration, Soochow University

摘要：需求預測對報童問題決策者十分重要，因為商品無法儲存至下期再販售。實務上，決策者常須在預算限制內選擇不同需求預測方法既而得到多個預測值，再加以組合。本文發展出一個存貨模型，幫助決策者以不等加權方式組合多個預測值來改善預測精確度。而最佳加權值是經由最小化組合預測變異數得來。本文發現兩不相關之預側值其加權值隨其變異數增加而遞減，因此當兩不相關之預側值在比較時，應選擇變異數較小之預側值優先組合。理論上，最佳組合可以完全搜尋演算法找到。但當被組合預測值個數過多時，完全搜尋演算法就變得沒有效率。於是本文提出向前搜尋演算法、向後搜尋演算法和相關搜尋演算法來解決。

關鍵詞：報童問題；需求預測；搜尋演算法

Abstract: Demand forecasting is important for the decision maker facing a newsboy problem as goods cannot be carried over to be sold in the following period. In this paper, we develop a model to assist the decision maker using an unequally weighted method in combining forecasts to improve forecast accuracy. The optimal weights are decided by minimizing the variance of combined

[1] Corresponding author: Department of Business Administration, Soochow University, Taipei City,

forecasts. We find that the optimal weights of uncorrelated forecasts decrease with their variances. When two uncorrelated forecasts are considered, one should select the forecast with smaller variance to combine with current forecasts in hand. Theoretically, the best combination of forecasts can be found by a complete search algorithm. We also propose three algorithms: a forward algorithm, a backward algorithm, and a correlated search algorithm to save computational time when the number of forecasts to be considered is large.

Keywords: Newsboy problem; Demand forecasting; Search algorithm

1. Introduction

In practice, forecasts of demand drive business planning, which involves tasks such as planning inventory and workforce levels, planning purchasing and production, budgeting, and scheduling. Thus, forecasting accuracy is one of the important factors that affect the effectiveness of business planning. Empirical studies show that forecasting accuracy is usually improved when forecasts are combined (Chan *et al.*, 1999). The newsboy problem can be used to handle the ordering of perishable items or style goods. Demand forecasting is important for a newsboy problem because the shelf life of goods in the problem is limited. Combining forecasts is seldom discussed in the papers regarding the newsboy problem. But papers in other fields have studied the combination of forecasts. For a more extensive review of the literature, please refer to Clemen (1989) and de Menezes *et al.* (2000).

Why should we combine forecasts? Clemen (1989) points out that if one can not recognize the underlying generating process of demand, it is better to combine forecasts from different forecasting methods that are able to capture different aspects of the information. How should forecasts be combined? Makridakis and Winkler (1983) find that the equally weighted method works well empirically, relative to the unequally weighted method. The popular approach to the unequally

Taiwan, E-mail: cmlee@scu.edu.tw

weighted method is to obtain the optimal weights by minimizing the mean squared forecast errors subject to the constraint that the weights sum to 1. Freeling (1981) shows that the weights will be larger for more accurate and less correlated forecasts; however, if the correlations between forecasts are strong and positive, the weights may be negative. Newbold and Granger (1974, 1984) note that in practice, users may find it expensive or impossible to obtain the covariance matrix of errors, and that the matrix is rarely stable over time. Bordley (1982, 1986) suggests a Bayesian approach to combine forecasts, and shows that under a normality assumption of forecast errors the optimal combination is a linear average of the forecasts, although an intercept is needed. Bates and Granger (1969) propose assigning the most weight to the model that has recently performed best. Bunn (1975) suggests an approach to assign weights that are proportional to the number of times that the model of interest has outperformed all other models to date. Weights of forecasts should be updated over time. Armstrong (2001) suggests updating the weights if evidence is strong. Winkler and Clemen (1992) develop graphs and sampling distributions for the weights. Deutsch *et al.* (1994) propose a method with changing weights that are derived from switching regression models or from smooth transition regression models. Chan *et al.* (2004) use cumulative sum (CUSUM) techniques to update the weights. Regarding the best number of forecasts to be combined, Makridakis and Winkler (1983) report that the accuracy of combined forecasts increases as more forecasts are combined; Bopp (1985) further reports that the accuracy tends to level off. Armstrong (2001) suggests combining at least five forecasts when possible.

The remainder of the paper is organized as follows. Section 2 proposes our model. Section 3 suggests rules and algorithms to find the optimal or near-optimal combination of forecasts and save computation time. Section 4 reports the results of numerical analysis. Section 5 concludes the paper.

2. The Model

Consider a newsboy problem. There are two stages in the decision process for each period. At the first stage, the decision maker has prior information of demand from historical data, but the decision maker can choose to buy other sources of information within a budget amount. After selection, forecasts from the selected sources are combined by using an unequally weighted method. Then, the combined forecast is used to update the prior demand to obtain the posterior demand. At the second stage, the decision maker decides the order quantity based on the posterior demand.

Let X be a random variable that represents the demand for the period. Assume that X follows $N(\theta, \tau^2)$ and that $h(x)$ is its density function. The distribution of X is considered as the prior information. At the first stage, the actual demand x is unknown. However, the decision maker may use expense-incurring information sources to estimate x; e.g., outside experts could be hired to provide forecasts that may be able to capture different aspects of the information (Clemen, 1989). Let $Y_i|x$ denote the conditional estimator of x from the *ith* source, whose distribution and forecast errors (or variance) can be obtained from past records of $Y_i|x$. In practice, outside experts are willing to provide the records of their past forecasts, or the decision maker can keep track of the performance of information sources. Then, $Y_i|x$ is assumed to follow $N(x, s_i^2)$, and also assumed to be an unbiased estimator of x. For a biased forecast, it should be obvious that a bias which is known will always be removed by offsetting. Therefore, we do not discuss the case of biased forecasts in this paper. Suppose there are n such sources available to choose from, thus $i = 1, 2,\ldots, n$. Let A be the set of selected information sources and b_i be a binary value. If $i \in A$, the *ith* source is selected and $b_i = 1$; otherwise, the *ith* source is not selected and $b_i = 0$. After selection, forecasts from the selected sources are obtained. Then, the decision maker is assumed to use an unequally weighted method to combine forecasts. Let λ_i be the weight of $Y_i|x$ and $Y|x$ be the combined forecast. Then,

$$Y|x = \sum_{i=1}^{n} \lambda_i b_i Y_i|x, \ \sum_{i=1}^{n} \lambda_i = 1, b_i = 0 \text{ or } 1, \text{if } b_i = 1, \lambda_i \neq 0; \text{if } b_i = 0, \lambda_i = 0.$$

It is clear that $Y|x$ follows $N(x, s^2)$, and the decision maker is assumed to decide the optimal weights of forecasts by minimizing the variance of combined forecast s^2:

$$\min_{\lambda_1,\dots,\lambda_n} s^2 = \min_{\lambda_1,\dots,\lambda_n} \{ \sum_{i=1}^{n} \lambda_i^2 b_i^2 s_i^2 + 2\sum_{i<j} \lambda_i \lambda_j b_i b_j COV(Y_i|x, Y_j|x) \},$$

$$\text{St. } \sum_{i=1}^{n} \lambda_i = 1, \ b_i = 0 \text{ or } 1, \ \text{if } b_i = 1, \lambda_i \neq 0; \text{if } b_i = 0, \lambda_i = 0. \qquad (1)$$

Let $\mathbf{1}^T = (1, 1, \dots, 1)$ and $\mathbf{\Sigma}$ be the covariance matrix of the forecasts. Solving (1) by Lagrangian method, we obtain the vector of the optimal weights $\mathbf{\Lambda}^{*T} = (\lambda^*_1, \lambda^*_2, \dots, \lambda^*_n)$, the optimal Lagrangian multiplier β^*, and the optimal s^2:

$$\mathbf{\Lambda}^* = \mathbf{\Sigma}^{-1}\mathbf{1}/(\mathbf{1}^T \mathbf{\Sigma}^{-1}\mathbf{1}), \qquad (2)$$

$$\beta^* = 2/(\mathbf{1}^T \mathbf{\Sigma}^{-1}\mathbf{1}), \qquad (3)$$

$$s^2 = 1/(\mathbf{1}^T \mathbf{\Sigma}^{-1}\mathbf{1}). \qquad (4)$$

The optimal weights in (2) are the same as those obtained by minimizing the mean squared forecast errors (Bates and Granger, 1969). In practice, models that assume independence between the individual forecast perform considerably better than those that attempt to consider correlation (Newbold and Granger, 1974, 1984). Besides, outside experts seldom provide the information of correlations among their forecasts. Thus, we assume that $Y_i|x$, $i = 1, 2, \dots, n$ are independent, i.e., $s^2 = \sum_{i=1}^{n} \lambda_i^2 b_i^2 s_i^2$. But, the cases of correlated forecasts are discussed in Sections 3.2 and 4.2.

Theorem 1. $\lambda_j^* = \dfrac{\frac{1}{s_j^2}}{\sum_{i\in A}\frac{1}{s_i^2}}, j\in A,\ \beta^* = \dfrac{2}{\sum_{i\in A}\frac{1}{s_i^2}}$, *and* $s^2 = \dfrac{1}{\sum_{i\in A}\frac{1}{s_i^2}}$.

Proof: λ^*_j, β^*, and s^2 are obtained by simplifying (2), (3), and (4), respectively.

From Theorem 1, forecasts with greater variances are given lower weights (see Freeling, 1981).

Lemma 1. *When one of two forecasts is considered to be included into A, one should include the forecast such that its variance is smaller, so that the decrease in s^2 is more significant.*

Proof: Lemma 1 is proved from Theorem 1.

Lemma 2. *s^2 decreases and tends to level off as more forecasts are combined.*

Proof: Lemma 2 is true from Theorem 1.

Let $g(y|x)$ be the density function of $Y|x$. Then, the density function $u(y)$ of Y is:

$$u(y) = \int_0^\infty g(y|x)h(x)dx = \frac{1}{\sqrt{2\pi}\sqrt{s^2+\tau^2}}\exp\{-(y-\theta)^2/2(s^2+\tau^2)\}. \quad (5)$$

The unconditional Y follows $N(\theta,\ s^2 + \tau^2)$. Next, the conditional $X|y$ is regarded as the posterior demand, and its density function $f(x|y)$ is obtained as follows.

$$f(x|y) = \frac{g(y|x)}{u(y)} = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-(x-\mu(y))^2}{2\sigma^2}\right), \quad (6)$$

$$\mu(y) = E(X|y) = (\tau^2 y + s^2\theta)/(\tau^2 + s^2), \quad (7)$$

$$\sigma^2 = Var(X|y) = \tau^2 s^2/(\tau^2 + s^2), \quad (8)$$

$$s^2 = \sum_{i=1}^{n}\lambda_i^2 b_i^2 s_i^2,\ \sum_{i=1}^{n}\lambda_i = 1, b_i = 0 \text{ or } 1. \quad (9)$$

From (7), the posterior mean $\mu(y)$ is a weighted average of the prior mean θ and the forecast y. Moreover, the weights are disproportionate to their variances.

Lemma 3. $\lim_{s^2\to\infty} \mu(y) = \theta$ *and* $\lim_{s^2\to\infty} \sigma^2 = \tau^2$.

Proof: This is easily shown from (7) and (8).

From Lemma 3, if s^2 approaches infinity, no forecast is selected and only the prior information is used; the posterior demand $X|y$ is reduced to the prior demand X. This may happen when the costs of sources are too expensive or uncertain.

Lemma 4. *σ^2 decreases with s^2 and the optimal weights in (2) also minimize σ^2.*

Proof: This is easily shown from (8), (1), (2), and Theorem 1.

Next consider the decision at the second stage. Let Q be the order quantity. Define $H(Q) = 0$ if $Q = 0$; otherwise, $H(Q) = 1$. Let $E(TC_o)$ be the expected total cost before ordering, and c_s, c_o, and c_u be the fixed ordering cost, unit overage cost, and unit underage cost, respectively. The decision maker's objective function is:

$$\min_{Q\geq 0} E(TC_o(Q)) = c_s H(Q) + c_u \int_Q^\infty (x-Q) f(x|y)dx + c_o \int_0^Q (Q-x) f(x|y)dx \,. \quad (10)$$

Equation (10) is a newsboy problem with fixed ordering cost. Let Q_1 be the optimal order quantity when fixed ordering cost is ignored. Then,

$\int_0^{Q_1} f(x|y)dx = c_u/(c_u+c_o) = \Phi(k_1)$, $k_1 = \Phi^{-1}(c_u/(c_u+c_o))$.

$Q_1 = \max\,\{\mu(y) + k_1\sigma, 0\}$.

(11) Let Q^* be the optimal order quantity. Then,

$$Q^* = \begin{cases} 0, & E(TC_o(0)) \leq E(TC_o(Q_1)) \\ Q_1, & E(TC_o(0)) > E(TC_o(Q_1)) \end{cases} \,. \quad (12)$$

Next, from (11), a condition for $Q^* > 0$ is:

$$\mu(y) + \sigma k_1 > 0. \quad (13)$$

Inserting (7) and (8) into (13), we have

$$y > -s^2\{k_1/\sigma + \theta/\tau^2\} = r. \tag{14}$$

Let Z denote $N(0, 1)$, and ϕ and Φ be its density function and distribution function, respectively. Let W follow $N(\gamma, \kappa^2)$ and $v(w)$ be its density function. We obtain the following formulas (Silver *et al.*, 1998).

$$\int_Q^\infty wv(w)dw = \tau\int_k^\infty (z-k)\phi(z)dz + Q\int_k^\infty \phi(z)dz = \tau\int_k^\infty z\phi(z)dz + \theta\int_k^\infty \phi(z)dz. \tag{15}$$

$$\int_k^\infty z\phi(z)dz = \phi(k), \quad k = (Q-\gamma)/\kappa.$$

(16)

Then, (10) can be simplified by (15) and (16) as follows.

$$E(TC_o) = \begin{cases} c_u\mu(y), & Q^* = 0 \\ c_s + (c_u + c_o)\sigma\phi(k_1), & Q^* > 0 \end{cases}. \tag{17}$$

From (12), for $Q^* > 0$, $E(TC_o(Q_1)) < E(TC_o(0))$. Therefore, from (17) we have

$$y > s^2\{c_s + (c_u + c_o)\sigma\phi(k_1)\}/c_u\sigma^2 - \theta s^2/\tau^2 = t. \tag{18}$$

Lemma 5. *There exists a threshold value of forecast y^* = max{r, t, 0}.*

Proof: Since $y > 0$, this is proved from (14) and (18).

From Lemma 5, if y, the value of the combined forecast $Y|x$, exceeds the threshold value y^*, an order will be issued; otherwise, nothing should be ordered.

Now, consider the decision at the first stage. At the first stage, the decision maker decides which information sources should be selected. The results of selection can be classified into two situations: A is empty and only prior information is used from Lemma 3; and A is not empty. After selection, forecasts are produced and combined to update the prior demand to obtain the posterior demand. The combined forecast has two effects on the ordering decision. One effect is that an order is issued if the combined forecast y is larger than the threshold value y^*. Let c_i be the cost of the ith forecast. The total cost when $Q^* > 0$ is as follows from Lemma 5 and (17):

$$TC_1 = (c_u + c_o)\sigma\phi(k_1) + c_s + \sum_{i=1}^{n} c_i b_i, \qquad y > y^*,\ b_i = 0 \text{ or } 1, \text{ and } A \neq \phi.$$

The other effect is that an order is not issued if the combined forecast does not exceed the threshold value. From Lemma 5 and (17), the total cost when $Q^* =$ 0 is:

$$TC_2 = c_u\mu(y) + \sum_{i=1}^{n} c_i b_i, \qquad y \leq y^*,\ b_i = 0 \text{ or } 1 \text{ and } A = \phi.$$

From (5), Y is $N(\theta, s^2 + \tau^2)$ and $u(y)$ is its density function. Let $E(TC_f)$ be the expected total cost before forecasting. Then, the objective function at the first stage is:

$$\min_{b_1,\dots,b_n} E(TC_f) = \begin{cases} \int_{y^*}^{\infty} TC_1 u(y)dy + \int_0^{y^*} TC_2 u(y)dy, & A \neq \phi. \\ c_s H(Q) + c_u \int_Q^{\infty} (x-Q)h(x)dx + c_o \int_0^Q (Q-x)h(x)dx, & A = \phi. \end{cases}$$

$$\text{St.} \quad \sum_{i=1}^{n} c_i b_i \leq b, b_i = 0 \text{ or } 1. \tag{19}$$

In (19), b is the budget for buying demand information. Equation (19), where A is not empty, is simplified by (15) and (16) as follows:

$$E(TC_f) = \{-c_u\theta + (c_u + c_o)\sigma\phi(k_1) + c_s\} \int_{k_2}^{\infty} \phi(z)dz - c_u\phi(k_2)\tau^2 / \sqrt{s^2+\tau^2}$$

$$+ c_u\theta + \sum_{i=1}^{n} b_i c_i, \qquad k_2 = (y^* - \theta)/\sqrt{s^2+\tau^2},$$

$$\text{St.} \quad \sum_{i=1}^{n} c_i b_i \leq b, b_i = 0 \text{ or } 1. \tag{20}$$

Equation (19), where A is empty, is simplified by (15) and (16) as follows:

$$E(TC_f) = \begin{cases} c_u\theta, & Q^* = 0 \\ c_s + (c_u + c_o)\tau\phi(k_1), & Q^* > 0 \end{cases}. \tag{21}$$

When the number of forecasts that could be chosen is small, (19) can be solved by the method of complete search and the number of evaluated combinations is 2^n. Next, we propose other search algorithms to save computational time when the number of forecasts is large.

3. The Heuristic Rules and Search Algorithms

3.1 Uncorrelated Forecasts

It is logical to infer from Lemma 1 that more accurate but less costly forecasts are preferred when selecting forecasts. Thus, we have the following heuristic rules.

Rule 1. *If $c_i \le c_j$ and $s_i \le s_j$, the ith forecast is no worse than the jth forecast and has a higher or equal priority to be included in A (the set of selected forecasts), or a lower or equal priority to be excluded from A.*

Rule 2. *If $c_1 \le c_2 \le ... \le c_n$ and $s_1 \le s_2 \le ... \le s_n$, the order of inclusion in A is 1, 2,..., n or the order of exclusion from A is n, n – 1,...,1.*

From Lemmas 2 and 4, adding more forecasts results in a more accurate $X|y$, but the accuracy tends to level off (Makridakis and Winkler, 1983, and Bopp, 1985). The cost of forecasts may finally outweigh the benefits of accuracy. Thus, we have

Rule 3. *Without considering the budget limit, when the forecasts are included one by one into A by the order decided by Rule 2, $E(TC_f)$ will decrease and then increase.*

Rule 4. *Without considering the budget limit, when the forecasts are excluded one by one from A by the order decided by Rule 2, $E(TC_f)$ will decrease and then increase.*

Rules 3 and 4 imply that $E(TC_f)$ is a concave upward function of the number of forecasts if forecasts are included by the "true" order of inclusion or excluded by the "true" order of exclusion. It is hard to prove these rules but they are discussed in the numerical analysis section. Since Rules 1 and 2 may be insufficient to decide the priorities for all forecasts, we define a cost-deviation index of the *ith* forecast as $c_i s_i$ to assist in deciding the order of inclusion or exclusion. The forecast with high cost-deviation index is considered to be less likely to be included or more likely to be excluded since its cost or standard deviation is high.

Rule 5. *The forecast with the lower value of cost-deviation index has a higher priority to be included in A or a lower priority to be excluded from A.*

Note that Rule 5 is also a heuristic rule based on Lemma 1. Next, we develop a forward algorithm and a backward algorithm by the concepts in Lemmas 1 and 2, and Rules 3, 4, and 5, to find the optimal or near-optimal combination and save computational time. For the two algorithms, the maximal number of evaluated combinations is $n + 1$, where n is the number of information sources. Comparing the forward and backward algorithms with the complete search algorithm, when $n = 5$ (Armstrong, 2001, suggests using five or more forecasts), the saving in computation time is at least 84%.

The forward algorithm assumes that A, the set containing forecasts to be combined, is empty then the forward algorithm begins to include forecasts one by one into A till the budget is violated or the expected cost increases. The backward algorithm assumes that all forecasts are already included in A then the backward algorithm begins to exclude forecasts one by one from A till the expected cost increases and budget is not violated.

Forward Algorithm

Step 1. Start from $A = \varnothing$ *and calculate* $E(TC_f)_A$ *by (21).*

Step 2. Determine the priority of inclusion for all forecasts by Rules 1 and 5.

Step 3. Include the current highest priority forecast into A. For the current A, check whether the budget constraint is violated. If yes, exclude the newly included forecast from A and go to Step 6.

Step 4. Calculate $E(TC_f)_A$ by (20). Check whether $E(TC_f)_A$ increases. If yes, exclude the forecast in Step 3 from A and go to Step 6; otherwise, go to the next step.

Step 5. If all forecasts are in A, go to Step 6; otherwise, let the forecast that has the next-highest priority be the current highest and go to Step 3.

Step 6. The current A is the solution set.

Backward Algorithm

Step 1. Start from A, including all forecasts. Calculate $E(TC_f)_A$ by (20).

Step 2. Determine the priority of exclusion for all forecasts by Rules 1 and 5.

Step 3. For the current A, check whether the budget constraint is satisfied. If yes, Feasible = Yes; otherwise, Feasible = No.

Step 4. Exclude the current highest-priority forecast from A. If A = $\varnothing$, calculate $E(TC_f)_A$ by (21); otherwise, calculate $E(TC_f)_A$ by (20).

Step 5. If Feasible = Yes and $E(TC_f)_A$ increases, add the excluded forecast in Step 4 into A and go to Step 7; otherwise, go to the next step.

Step 6. If A = $\varnothing$, go to Step 7; otherwise, let the forecast with the next-highest priority be the current highest and go to Step 3.

Step 7. The current A is the solution set.

3.2 Correlated Forecasts

For correlated forecasts, weights and s^2 are calculated by (2) and (4), respectively. Let $A_{k,m}$ be the *m*th combination set when *k* correlated forecasts are selected. Define $c_{k,m}$ to be its corresponding total cost of forecasts and $s_{k,m}$ to be the corresponding combined conditional standard deviation. For example, there are five correlated forecasts to be considered, then $A_{2,3} = \{1, 4\}$, $s_{2,3} = (\lambda_1^2 s_1^2 + \lambda_4^2 s_4^2 + 2\lambda_1\lambda_4 COV(Y_1|x, Y_4|x))^{1/2}$ and $c_{2,3} = c_1 + c_4$.

Rule 6. *If $c_{k,m} \le c_{k,m'}$, and $s_{k,m} \le s_{k,m'}$, $A_{k,m}$ is no worse than $A_{k,m'}$ and has a higher or equal priority to be selected.*

We define a combined cost-deviation index of $A_{k,m}$ as $c_{k,m}s_{k,m}$. Then, we have

Rule 7. *The set with the lower value of index has a higher priority to be selected.*

Rules 6 and 7 are also heuristic rules. The following algorithm is proposed to find the optimal or near-optimal set. The expected number of evaluated combinations is $n + 1$.

Correlated Search Algorithm

Step 1. Let k = 0 and Tempcost = ∞.

Step 2. Find all sets containing k forecasts.

Step 3. For those sets in Step 2 that do not violate the budget constraint, calculate their indexes. Find the selection set with the highest priority by Rule 7, and calculate its $E(TC_f)$ *using (20) or (21).*

Step 4. Let Tempcost(k) be the $E(TC_f)$ *in Step 3 and Tempset(k) be the corresponding selection set.*

Step 5. If Tempcost(k) $\leq$ *Tempcost, then Tempcost = Tempcost(k) and Tempset =Tempset(k). Otherwise, nothing is done.*

Step 6. If k = n, go to the next step. Otherwise, k = k + 1. Go to Step 2.

Step 7. Tempcost is the optimal cost and Tempset is the optimal solution set.

4. Numerical Analysis

Set the values of some model parameters in Section 4 to be as follows.

X the prior demand follows $N(\theta, \tau^2)$, $\theta = 5{,}000$, and $\tau^2 = 1{,}500^2$.

c_u unit underage cost is \$2.4.

c_o unit overage cost is \$2.

c_s fixed ordering cost is \$4,500.

4.1 Uncorrelated Forecasts

We first investigate the validity of Rules 1 and 5. Suppose the decision maker has to choose one of two forecasts. Consider the following three cases. In Case 1, the variances of the two forecasts are $s_1^2 = s_2^2 = 1{,}400^2$, the costs of the two forecasts are $c_2 = 200$, and c_1 changes from 150 to 550. In Case 2, $c_1 = c_2 =$

200, s_2^2 = 1,600^2, and s_1 changes from 1,200 to 2,000. The results of computer output are in Tables 1 and 2. We find that Rules 1 and 5 are useful since correct decisions are made.

Table 1

The Validity of Rules 1 and 5 for Uncorrelated Forecasts in Case 1

C_1	$E(TC_f)_{\{1\}}$	$E(TC_f)_{\{2\}}$	C_1S_1	C_2S_2	Best Set	Rule 1	Rule 5
150	6,421	6,671	210,000	560,000	{1}	Correct	Correct
250	6,521	6,671	350,000	560,000	{1}	Correct	Correct
350	6,721	6,671	490,000	560,000	{1}	Correct	Correct
450	6,821	6,671	630,000	560,000	{2}	Correct	Correct
550	6,921	6,671	770,000	560,000	{2}	Correct	Correct

Table 2

The Validity of Rules 1 and 5 for Uncorrelated Forecasts in Case 2

S_1	$E(TC_f)_{\{1\}}$	$E(TC_f)_{\{2\}}$	C_1S_1	C_2S_2	Best Set	Rule 1	Rule 5
1,200	6,315	6,599	240,000	320,000	{1}	Correct	Correct
1,400	6,471	6,599	280,000	320,000	{1}	Correct	Correct
1,600	6,599	6,599	320,000	320,000	{1} or {2}	Correct	Correct
1,800	6,703	6,599	360,000	320,000	{2}	Correct	Correct
2,000	6,788	6,599	400,000	320,000	{2}	Correct	Correct

Next, in Case 3, c_1 = 200, c_2 = 400, s_2^2 = 1,400^2, and s_1 changes from 1,200 to 3,600. The results are in Table 3. We find that Rule 1 is useful when it is applicable. When $s_1 \geq$ 3,000, Rule 1 can not be applied but correct decisions are made by using Rule 5. Thus, Rule 5 is useful. However, when indices are close to each other (or c_1s_1 is between 360,000 and 480,000), wrong decisions are made by using Rule 5 and the cost error is within 3.69%.

Table 3

The Validity of Rules 1 and 5 for Uncorrelated Forecasts in Case 3

S_1	$E(TC_f)_{\{1\}}$	$E(TC_f)_{\{2\}}$	C_1S_1	C_2S_2	Best Set	Rule 1	Rule 5
1200	6,315	6,671	240,000	560,000	{1}	Correct	Correct
1800	6,703	6,671	360,000	560,000	{2}	N.A.*	Wrong
2400	6,917	6,671	480,000	560,000	{2}	N.A.	Wrong
3000	7,039	6,671	600,000	560,000	{2}	N.A.	Correct
3600	7,114	6,671	720,000	560,000	{2}	N.A.	Correct

Note : *: Not applicable.

In Case 4, five uncorrelated forecasts are considered to be combined. Set the new values of the model parameters to be:

$Y_i|x$ the *ith* forecast follows $N(x, s_i^2)$, $i = 1, 2, 3, 4$, and 5.

s_i^2 $s_1^2 = 1{,}400^2$, $s_2^2 = 1{,}400^2$, $s_3^2 = 1{,}600^2$, $s_4^2 = 1{,}500^2$, and $s_5^2 = 1{,}300^2$.

c_i the cost of $Y_i|x$, c_1=\$200, c_2=\$400, c_3 =\$200, c_4=\$600, and c_5=\$250.

b the budget amount is \$1,500.

The order of inclusion can not be determined by Rule 1. Then, in computing the cost-deviation index $c_i s_i$, by Rule 5, we find that the order of inclusion is 1, 3, 5, 2, and 4 or the order of exclusion is 4, 2, 5, 3, and 1. The results are in Tables 4 and 5.

Table 4

The Expected Costs for Sets of Uncorrelated Forecasts in Case 4

Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$
∅	7,116	{1,3}	6,381	{3,5}	6,386	{1,4,5}	6,757	{1,2,4,5}	7,022
{1}	6,471	{1,4}	6,751	{4,5}	6,759	{2,3,4}	6,976	{1,3,4,5}	6,850
{2}	6,671	{1,5}	6,327	{1,2,3}	6,552	{2,3,5}	6,575	{2,3,4,5}	7,050
{3}	6,599	{2,3}	6,581	{1,2,4}	6,934	{2,4,5}	6,957	{1,2,3,4,5}$^+$	7,144
{4}	6,938	{2,4}	6,951	{1,2,5}	6,537	{3,4,5}	6,797		
{5}	6,447	{2,5}	6,527	{1,3,4}	6,776	{1,2,3,4}	7,020		
{1,2}	6,517	{3,4}	6,820	{1,3,5}	6,375	{1,2,3,5}	6,634		

Note :+: Violate budget constraint.

Table 5

The Found sets for Various Numbers of Forecasts in Case 4

No. of Forecasts	The Complete Search Algorithm	The Forward Algorithm	The Backward Algorithm
0	∅	∅	N.A
1	{5}	{1}	N.A
2	{1,5}*	{1,3}	{1,3}
3	{1,3,5}	{1,3,5}$^\#$	{1,3,5}$^\#$
4	{1,2,3,5}	{1,2,3,5}	{1,2,3,5}
5	{1,2,3,4,5}$^+$	N.A.	{1,2,3,4,5}$^+$

Note: $^{:\#}$: Also found as the optimal set by a different algorithm

From Tables 4 and 5, the optimal set is {1, 5}, and the expenditure on forecasting is \$450 (=$c_1 + c_5$). In using the complete search algorithm, the number of evaluated sets is 32 (=2^5). In using the forward and backward algorithm, the number of evaluated sets is 5 and 4, respectively, an 84% reduction compared to 32 evaluated sets for the complete search algorithm; the near optimal set {1, 3, 5} is found and the cost error is \$48, which is about 0.76% of the expected optimal cost (\$6,327). Thus, use of the forward and backward algorithms may considerably reduce computation burden within reasonable cost errors. We find that the differences in cost-deviation indices among Forecasts 1, 3, and 5 are relatively small compared to those involving other forecasts. Thus, the forward and backward algorithms fail to find the optimal set. Then, the following rule can be used to find the optimal set.

Rule 8. *Revise the order of inclusion without violating Rule 1. Find out all possible revised orders. Then, apply forward or backward algorithms for each revised order to find the optimal sets and their corresponding optimal expected costs. Then, compare these costs with the cost of the optimal set of the original order to find any cost improvement.*

The effectiveness of Rule 8 depends on the number of possible revised orders. In this case, 1, 5, 3, 2, and 4 is a possible revised order. From the results of Tables 4 and 5, the optimal set in Case 4 can be found by Rule 8.

4.2 Correlated forecasts

Suppose two correlated forecasts are considered to be combined. In Case 5, $c_1 = c_2 = 200$, $s_1^2 = 1{,}400^2$ and $s_2^2 = 1{,}600^2$, and covariance changes from -1,792,000 to 1,792,000. Since $-1 \le \rho \le 1$, $-\sigma_x\sigma_y \le COV(x, y) \le \sigma_x\sigma_y$. The results are in Table 6. From Table 6, when the coefficient of correlation is positive, the best set is {1}; otherwise, the best set is {1, 2}, and the weight of Forecast 2, λ_2, decreases with covariance. It seems better to combine negatively correlated forecasts because overestimated forecasts may be traded off by other underestimated forecasts. In other words, it is less useful to combine "redundant" forecasts, i.e. highly positively correlated forecasts.

Table 6

Comparison Between Positively and Negatively Correlated Forecasts in Case 5

$COV_{1,2}$	ρ	$E(TC_f)_{\{1\}}$	$E(TC_f)_{\{2\}}$	$E(TC_f)_{\{1,2\}}$	Best Set	λ_1	λ_2
-1,792,000	-0.92	6,471	6,599	5,653	{1,2}	0.533	0.467
-896,000	-0.46	6,471	6,599	6,123	{1,2}	0.542	0.458
0	0	6,471	6,599	6,381	{1,2}	0.555	0.445
896,000	0.46	6,471	6,599	6,551	{1}	0.583	0.417
1,792,000	0.92	6,471	6,599	6,663	{1}	0.664	0.336

Table 7

The Validity of Rules 6 and 7 for Correlated Forecasts in Case 6

C_3	$E(TC_f)_{\{1,2\}}$	$E(TC_f)_{\{1,3\}}$	$C_1 + C_3$	Index of {1,3}	Best Set	Rule 6	Rule 7
50	6,027	6,398	300	353,834	{1,2}	N.A.	Wrong
250	6,027	6,598	500	589,724	{1,2}	N.A.	Correct
450	6,027	6,798	700	825,613	{1,2}	Correct	Correct
650	6,027	6,998	900	1061,503	{1,2}	Correct	Correct
850	6,027	7,198	1,100	1297,393	{1,2}	Correct	Correct

Next, we study the validity of Rules 6 and 7. In Case 6, three forecasts are considered and suppose that Forecast 1 is already selected. Furthermore, $c_1 = 250$, $c_2 = 400$, c_3 changes from 50 to 850, $s_1^2 = 1,300^2$, $s_2^2 = 1,400^2$, $s_3^2 = 1,600^2$, the covariance of Forecasts 1 and 2, $COV_{1,2}$, is -120,000, and $COV_{1,3} = 80,000$. Note that in this case, $c_1 + c_2 = 650$, the combined standard deviation of {1, 2} is 556 (rounding to integer), the cost-deviation index of {1, 2} is 361,606, and the combined standard deviation of {1, 3} is 1,179. The other results are in Table 7. We find that Rule 6 is useful when it is applicable. When c_3 is about 250, Rule 6 can not be applied but correct decisions are made by using Rule 7. Thus, Rule 7 is useful. However, when c_1 is about 50, the combined cost-deviation indices are close to each other (one is 6,027 and the other is 6,398). Wrong decisions are made by using Rule 7 and the cost error is within 6.16%.

In Case 7, $COV_{1,3} = 20,000$, s_3 changes from 500 to 2,100, and the values of other parameters are the same as those in Case 6. Note that in Case 7, $c_1 + c_2 = 650$, $c_1 + c_3 = 450$, the combined standard deviation of {1, 2} is 556, and the cost-deviation index of {1, 2} is 361,606. The other results are in Table 8. From

Table 8, Rule 6 is useful when it is applicable. When $1{,}300 \le s_1$, Rule 6 can not be applied but correct decisions are made by using Rule 7. Thus, Rule 7 is useful. However, when s_1 is about 900, the combined cost-deviation indices are close to each other (one is 6,027 and the other is 6,147). Wrong decisions are made by using Rule 7, and the cost error is within 1.99%.

Table 8

The Validity of Rules 6 and 7 for Correlated Forecasts in Case 7

S_3	$E(TC_f)_{\{1,2\}}$	$E(TC_f)_{\{1,3\}}$	Std of {1,3}^	Index of {1,3}	Best Set	Rule 6	Rule 7
500	6,027	5,742	498	224,268	{1,3}	Correct	Correct
900	6,027	6,147	795	357,972	{1,2}	N.A.	Wrong
1,300	6,027	6,348	972	437,450	{1,2}	N.A.	Correct
1,700	6,027	6,453	1,077	484,430	{1,2}	N.A.	Correct
2,100	6,027	6,517	1,145	515,224	{1,2}	N.A.	Correct

Note :^: Rounding to integer

Table 9

The Validity of Rules 6 and 7 for Correlated Forecasts in Case 8

$COV_{1,3}$	ρ	Std of {1,3}^	Index of {1,3}	$E(TC_f)_{\{1,3\}}$	Best Set	Rule 6	Rule 7
-1,664,000	-0.8	459	204,009	5,674	{1,3}	Correct	Correct
-832,000	-0.4	784	352,757	6,132	{1,2}	N.A.	Wrong
0	0	1,009	454,027	6,386	{1,2}	N.A.	Correct
832,000	0.4	1,185	533,460	6,553	{1,2}	N.A.	Correct
1,664,000	0.8	1,300	584,874	6,647	{1,2}	N.A.	Correct

In Case 8, $COV_{1,3}$ changes from -1,664,000 to 1,664,000, and the values of other parameters are the same as those in Case 6. Note that in Case 8, $c_1 + c_2 = 650$, $c_1 + c_3 = 450$, the combined standard deviation of {1, 2} is 556, and the cost-deviation index of {1, 2} is 361,606. The other results are in Table 9. We find that Rule 6 is useful when it is applicable. When $0 \le COV_{1,3}$, Rule 6 can not be applied but correct decisions are made by using Rule 7. Thus, Rule 7 is useful. However, when $COV_{1,3}$ is about -832,000, the combined cost-deviation indices are close to each other (one is 361,606 and the other is 352,757). Wrong decisions are made by using Rule 7 and the cost error is within 1.74%.

Next, consider Case 9. In this case, the setting and the values of model parameters are the same as those in Case 4. Besides, the covariance matrix is a positive semi-definite symmetric matrix as follows. The combined cost-deviation index $c_{k,m}s_{k,m}$ are computed by Rules 6 and 7. Applying the correlated search algorithm and the complete search algorithm, the respective results are in Tables 10 and 11.

$$\Sigma = \begin{bmatrix} 1960000 & -1000000 & 800000 & -700000 & 600000 \\ & 1960000 & -1300000 & 400000 & -1200000 \\ & & 2560000 & -900000 & 800000 \\ & & & 2250000 & -500000 \\ & & & & 1690000 \end{bmatrix}$$

Table 10

The Expected Costs for Sets of Correlated Forecasts

Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$	Set	$E(TC_f)$
∅	7,116	{1,3}	6,536	{3,5}	6,548	{1,4,5}	6,642	{1,2,4,5}	6,582
{1}	6,471	{1,4}	6,545	{4,5}	6,610	{2,3,4}	6,583	{1,3,4,5}	6,782
{2}	6,671	{1,5}	6,473	{1,2,3}	6,151	{2,3,5}	6,075	{2,3,4,5}	6,571
{3}	6,599	{2,3}	6,154	{1,2,4}	6,615	{2,4,5}	6,549	$\{1,2,3,4,5\}^{+}$	6,679
{4}	6,938	{2,4}	7,041	{1,2,5}	6,075	{3,4,5}	6,675		
{5}	6,447	{2,5}	6,027	{1,3,4}	6,602	{1,2,3,4}	6,592		
{1,2}	6,166	{3,4}	6,580	{1,3,5}	6,640	{1,2,3,5}	6,194		

Table 11

The Found Sets for Various Numbers of Correlated Forecasts

No. of Forecasts	The Complete Search Algorithm	The Correlated Algorithm
0	∅	∅
1	{5}	{1}
2	$\{2,5\}^{\#}$	$\{2,5\}^{\#}$
3	{1,2,5}	{1,2,5}
4	{1,2,3,5}	{1,2,3,5}
5	$\{1,2,3,4,5\}^{+}$	$\{1,2,3,4,5\}^{+}$

By the correlated search algorithm, the optimal set {2, 5} is found, and Forecasts 2 and 5 are negatively correlated. Thus, it is better to combine negatively correlated forecasts. However, if the near-optimal set is found, the following rule is useful.

Rule 9. *Revise the order of inclusion without violating Rule 6. Find out all* *possible revised orders. Then, apply forward or backward algorithms for each revised order to find the optimal sets and their corresponding optimal expected costs. Then, compare these costs with the cost of the optimal set of the original order to find any cost improvement.*

Finally, Figure 1 depicts $E(TC_f)$ when using the true order of inclusion in Cases 4 and 9. The expected costs decrease then increase (concave upward) in both cases. It shows that Rules 3, 4, 5, 6, and 7 are useful.

Figure 1

Expected cost and the number of combined forecasts

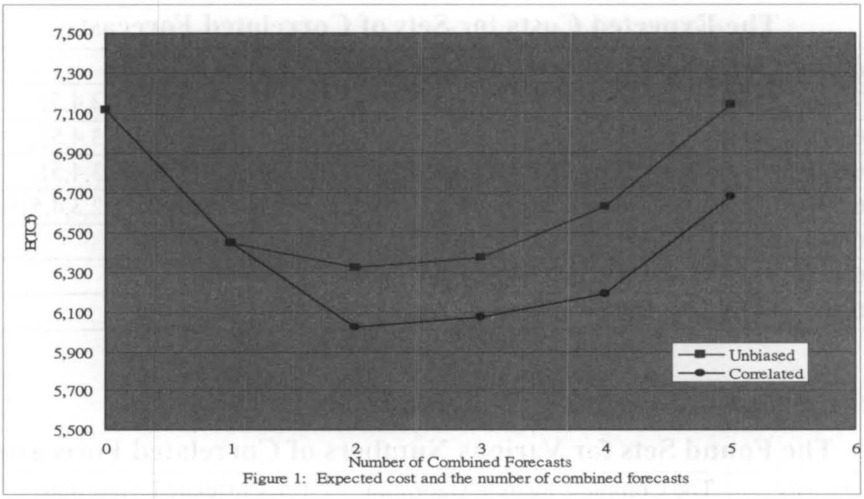

Figure 1: Expected cost and the number of combined forecasts

5. Conclusions

This paper develops a model for a decision maker facing a newsboy problem to combine demand forecasts. Consider a sequential decision process in a newsboy problem. The decision maker already has the prior information of demand from past records, but there are different sources of demand information that may be purchased. The decision maker needs to decide which sources to be purchased without violating the budget. After selection, forecasts from selected sources are combined to update the prior demand to obtain the posterior demand. The order quantity is then decided based on the posterior demand. The present results confirm some results of previous studies. Other important results are as follows.

(i). The optimal weights of uncorrelated forecasts decrease as their variances increase.

(ii). Without considering the costs of forecasts, when two uncorrelated forecasts are compared, we select the forecast with smaller variance to combine with current forecasts in hand.

(iii). It is better to combine negatively correlated forecasts.

(iv). When no forecast is selected, the decision maker uses the prior information only.

(v). When the number of uncorrelated forecasts is large, we suggest the use of the forward algorithm or the backward algorithm to find the optimal set or near-optimal set within a reasonable cost error and to reduce computation time; as for the correlated forecasts, we suggest the use of the correlated search algorithm. Use of the complete search algorithm to find the optimal set is suggested when the number of forecasts is small.

References

Armstrong, S. (2001), *Principles of Forecasting: A Handbook for Researcher and Practitioners*, Dordrecht, Netherlands: KAP.

Bates, J. M. and Granger, C. W. J. (1969), "The Combination of Forecasts," *Operational Research Quarterly*, 20(4), 451-468.

Bopp, A. (1985), "On Combining Forecasts: Some Extensions and Results," *Management Science*, 31(12), 1492-1498.

Bordley, R. F. (1982), "The Combination of Forecasts: a Bayesian Approach," *Journal of the Operational Research Society*, 33(2), 171-174.

Bordley, R. F. (1986), "Technical Note: Linear Combination of Forecasts with an Intercept: a Bayesian Approach," *Journal of Forecasting*, 5(4), 243-249.

Bunn, D. W. (1975), "A Bayesian Approach to the Linear Combination of Forecasts," *Operational Research Quarterly*, 26(2), 325-329.

Chan, C. K., Kingsman, B. G. and Wong, H. (1999), "The Value of Combining of Forecasts in Inventory Management: A Case Study in Banking," *European Journal of Operational Research*, 117(2), 199-210.

Chan, C. K., Kingsman, B. G. and Wong, H. (2004), "Determining When to Update the Weights in Combined Forecasts for Product Demand—An Application of the CUSUM Technique," *European Journal of Operational Research*, 153(3), 757-768.

Clemen, R. T. (1989), "Combining Forecasts: a Review and Annotated Bibliography," *International Journal of Forecasting*, 5(4), 559-583.

de Menezes, L. M., Bunn, D. W. and Taylor, J. W. (2000), " Review of Guidelines for the Use of Combined Forecasts," *European Journal of Operational Research*, 120(1), 190-204.

Deutsch, M., Granger, C. W. J. and Terasvirta, T. (1994), "The Combination of Forecasts Using Changing Weights," *International Journal of Forecasting*, 10(1), 47-57.

Freeling, A. N. S. (1981), "Reconciliation of Multiple Probability Assessments," *Organizational Behavior and Human Performance*, 28(3), 395-414.

Makridakis, S. and Winkler R. L. (1983), "Averages of Forecasts: Some Empirical Results," *Management Science*, 29(9), 987-996.

Newbold, P. and Granger C. W. J. (1974), "Experience with Forecasting Univariate Time Series and the Combination of Forecasts (with Discussion)," *Journal of the Royal Statistical Society Series A*, 137(2), 131-149.

Newbold, P. and Granger C.W.J. (1984), "Experience with Forecasting Univariate Time Series and the Combination of Forecasts," *Journal of Marketing Research*, 21(3), 290-297.

Silver, E. A., Pyke, D.F. and Peterson, R. (1998), *Inventory Management and Production Planning and Scheduling*, New York, NY: Wiley.

Winkler, R. L. and Clemen, R. T. (1992), "Sensitivity of Weights in Combining Forecasts," *Operations Research*, 40(3), 609-14.