Yugoslav Journal of Operations Research
28 (2018), Number 2, 265–274
DOI: https://doi.org/10.2298/YJOR171116015B

FINANCIAL PLANNING IN THE PUBLIC SECTOR AND THE NEWSVENDOR PROBLEM

Jack BRIMBERG
Mathematics and Computer Science department, The Royal Military College of Canada
Jack.Brimberg@rmc.ca

Abdalla MANSUR
Al Ain University of Science and Technology
Al Ain, Abu Dhabi United Arab Emirates
abdalla.mansur@aau.ac.ae

Ghaleb A. EL-REFAE
Al Ain University of Science and technology
Al Ain, Abu Dhabi United Arab Emirates
president@aau.ac.ae

Received: October 2017 / Accepted: April 2018

Abstract: The purpose of this paper is to study the phenomenon of lapsed funding that occurs within the public sector using the framework of the Newsvendor Problem. Under this framework, we observe that lapsed funding is a natural outcome of the planning system when the financial manager is trying to maximize the value added to his unit. Centralizing control of operating budgets of homogeneous units is shown to reduce lapsed funds and improve their overall performance.

Keywords: Newsvendor Problem, Lapsed Funds, Public Sector.

MSC: 91B02, 91B06, 97M10, 97M40.

1. INTRODUCTION

The Newsvendor Problem (NVP), originally known as the Newsboy Problem, is a widely used model in management decision processes. Typically, the vendor or decision maker wants to determine how much of a particular good to order when demand is uncertain and the good has a limited shelf-life. He must make this decision before demand is realized. Thus, the decision-maker must trade off the costs of having too much supply versus having too little; see e.g. [6, 7, 8] for good reviews of the NVP literature.

In this paper we review an interesting application of the NVP to defence spending, which also extends to financial planning in other public sector departments. For a detailed development of the mathematical model, see [3], and [4] for an extension to multiple decision points. One of the main problems facing public sector financial managers is the shelf-life of their operating funds. Typically, a government department will be allocated a fixed amount of funding for the year. In the absence of a system that allows carryovers, any operating funds not expended over the year may be lost to the department permanently. These are termed lapsed funds and in most cases, a manager's performance review will be adversely affected if his/her lapsed funding is too high.

One of the difficulties faced by public sector managers relates to the uncertain costs of the activities undertaken. This is particularly true in defence departments. The problem is to pick the right portfolio of high-value activities over the year so that lapsed funding is kept to a minimum. However, as argued here, the finite shelf-life of available dollars coupled with the necessary timing of decisions make this an analogous problem to the NVP. Thus, lapsed funds become an unavoidable by-product imposed by the constraints of the system.

Our research also has a practical implication for defence managers and the incentives these managers operate under. Namely, if managers are doing their jobs properly, there will be years when some funds will remain, and hence will lapse. That is, in a newsvendor environment, lapsed funds can actually be an indication of efficiency, and thus some reasonable percentage of lapsed funds should be viewed in a positive light. The paper consists of four parts. We first give a brief review of the Newsvendor Problem. Then we review how the management of defence operating funding has a structure equivalent to this problem. The pooling of homogeneous units at the planning level is subsequently shown to improve performance of the individual units. We conclude with a general discussion on the usefulness of the NVP framework within the context of financial planning in the public sector.

2. THE NEWSVENDOR MODEL

Consider this familiar scene: It is around suppertime, and you stop at the local grocery store on the way home from work to get a copy of your favorite newspaper. To your dismay, you find the shelf completely empty. Alternatively, it may be near midnight, just before closing time, when you remember you need the paper. To your surprise you find the same shelf still well-stocked with today's issue, and wonder what the fellow at the cash will do with those useless extra copies. These scenes characterize the fundamental nature of the classical newsvendor (or newsboy) problem. The newsvendor faces the same basic dilemma at the start of each day: how many copies of the newspaper should the vendor order to meet the uncertain demand for this particular product. If he or she orders an insufficient quantity of the newspaper, the store will run out of stock, and hence, lose the opportunity to make more profit (as well as anger some customers). But if he orders too much stock, he will be left with unsold copies of the paper at the end of the day. Since the shelf-life has expired by then, the net effect is that his income is reduced by the cost of these leftover copies less, perhaps, a small salvage value.

The problem may be formulated mathematically. Let D denote the daily demand for the particular newspaper. Assuming that the vendor has kept track of his daily sales, he would certainly have a good feel for the probability distribution of this random variable. Suppose the paper costs him c dollars per unit to procure, and he sells it at p dollars per unit. Unsold copies at the end of the day are returned to the newspaper company who reimburses the vendor at a salvage rate of c_s per unit. We may now apply a marginal analysis where D is treated as a continuous random variable:

- The expected daily marginal gain accruing to the vendor by increasing his order size from a given number of units x_0 to $x_0 + \delta x$ would be

$$E[G] = (p - c)P\{D > x_0\} - (c - c_s)P\{D \leq x_0\}$$

- As long as $E[G] \geq 0$, the rational vendor will increase the order size (x), since this will improve his average daily profit in the long run. Thus, the rational vendor will increase x until $E[G] = 0$. Letting $F(x)$ denote the cumulative distribution function of the daily demand D, we obtain:

$$E[G] = (p - c)(1 - F(x)) - (c - c_s)F(x) = 0.$$

Thus, the optimal order size x^* satisfies the following equation:

$$F(x^*) = \frac{p - c}{p - c_s} \tag{1}$$

Using the notation

$$r = \frac{p - c}{p - c_s}, \tag{2}$$

we have for the optimal order size x^*:

$$P(stockout) = P(D > x^*) = 1 - r, \tag{3}$$

$$P(no\ \ stockout) = P(D \leq x^*) = r. \tag{4}$$

Furthermore, the chance of the vendor getting it exactly right on any day, $P(D = x^*)$, will be very small ($\to 0$).

3. COMPARISON WITH THE FINANCIAL PLANNER

What do the newsvendor on the street corner and the financial planner in the public sector have in common? It turns out to be an interesting comparision! We have just seen that the news vendor must deal with uncertain demand, and he must make his decision on the number of papers to order at the start of the planning period, in his case, the day. The odds that the number of papers he orders will be equal to the demand in a given day are pretty slight, let's say zero. Thus, he will surely run out of stock by the end of the day or have unsold copies remaining on the shelf. Let's see now what the situation looks like for the financial planner at the start of the fiscal year.

The planner will undoubtedly have several different budgets to manage, but we are interested primarily in the *operating* budget. At the beginning of the year, the planner is given a set of activities that his unit would like to accomplish during the year. Some of these activities are critical to the success of the unit, for example, these might include training activities for staff, vital maintenance activities, and so on. In addition there would be nice-to-do, but non-critical activites such as optional upgrades of computer equipment, refurbishment of facilities, and the like. The set of activities are listed in order of priority so that those that are essential to the operations are at the top. The planner also receives a total budget for the year that his unit should not exceed.

To formalize the discussion, let the activities listed in order of decreasing priority be denoted by $A_1, A_2, .., A_M$. Each activity A_i has a budgeted cost b_i that also should not be exceeded. Let b_T denote the total budget allocated to the unit to meet its operational needs. It would be nice if this budget were large enough to fund all the listed activities, but typically it falls well short of the mark. Thus, the planner is only guaranteed funds to complete activities $A_1, A_2, ..., A_n$ where $n < M$. These are termed the *planned* (or *programmed*) activities. If the planner only selects these planned activities, he will find out later in the year, almost with certainty and when it may be too late to start up any substantial remaining activities, that there will be budget remaining unspent by year's end; i.e., his unit will be stuck with lapsed funds.

Let Q_i denote the actual funds that will be expended on activity $A_i, i = 1, ..., n$, over the course of the year. Since $Q_i \leq b_i$, the total expenditure at the end of the year for the *planned* activites will be:

$$Q = \sum_{i=1}^{n} Q_i \leq \sum_{i=1}^{n} b_i \leq b_T,$$

so that the unspent budget if no other activites are selected is given by:

$$U = b_T - Q \geq 0.$$

This uncertain value (or random variable) U is referred to as total slippage.

In order to use the total available budget as effectively as possible (as well as avoid the unpleasant situation of lapsed funding), the planner is allowed to schedule additional activities at the start of the year. These activities, $A_{n+1}, A_{n+2}, ..., A_{n+t}$, selected from the remaining ones in order of priority are termed *over-planned* (*over-programmed*) activities because funding for them is not guaranteed. The planner's basic decision at the start of the year thus boils down to determining what level of *over-programming* to commit to at this time, i.e., how far down the list to go. If he does not go down the list far enough, there will be lapsed funds; if he moves too far down, the available budget is exceeded, in which case some activities will have to be reduced or stopped (referred to as *off-ramping*). With this background, we are now ready to compare the planner's dilemma and the newsvendor's.

The key observation here is that the slippage U, which represents the unknown availability of over-programming dollars, may also be interpreted as a demand for these dollars. The planner does not know what the value of U will be when he decides how many of these *over-programming* dollars to procure. As in the case of the newsvendor, he probably has some idea of the probability distribution of U from his experience and any available historical data. In short the planner's decision reduces to the number of *over-programming* dollars to procure, which we denote by the variable x. If the value of x chosen turns out to be too low ($x < U$), there will be lapsed funds; if it is too high ($x > U$), the available budget is exceeded, and *off-ramping* must occur.

Just as with the newsvendor, the objective of the rational planner is to choose a value of x that maximizes value, although the measurement of value is not as straight forward in this case. Sidestepping this issue we let v_{OP} denote the value of each additional over-programming dollar that is utilized, and v_{OR} denote the lost value of one dollar of off-ramped activities. Since off-ramping by its nature also necessitates the closing down of some planned activities, as well as possible penalty costs, it follows that $v_{OP} < v_{OR}$. The net loss per dollar of off-ramping would thus equal $v_{OR} - v_{OP} > 0$. Using the same type of marginal analysis, and letting $F_U(u)$ denote the cumulative distribution function of U, we obtain at optimality

$$E[G] = v_{OP}(1 - F_U(x)) - (v_{OR} - v_{OP})F_U(x) = v_{OP} - v_{OR}F_U(x) = 0.$$

Thus, the optimal level of *over-programming* x^* must satisfy the relation,

$$F_U(x^*) = r, \tag{5}$$

where the ratio is now given by:

$$r = \frac{v_{OP}}{v_{OR}}. \tag{6}$$

Note that $r < 1$, since $v_{OP} < v_{OR}$.

As with the newsvendor, the financial planner will rarely get it right. Faced with uncertain demand for *over-programming* dollars and the imperative to decide at (or relatively near) the beginning of the year how many of these dollars to procure, the planner will inevitably be left with either an under-spent or over-spent budget towards the end of the year. The respective probabilities are:

$$P\{lapsed\ funds\} = P\{U > x^*\} = 1 - F_U(x^*) = 1 - r, \tag{7}$$

$$P\{off - ramping\} = P\{U < x^*\} = F_U(x^*) = r. \tag{8}$$

Typically, the number of *planned* activities (n) is quite large, so that under reasonable assumptions (such as statistical independence of activity costs, Q_i), it follows by Lyapunov's Central Limit Theorem [2] that total slippage U behaves, at least in approximation, as a normal random variable. Letting μ_U and σ_U denote, repectively, the mean and the standard deviation of U, and using the conventional notation $\Phi(z)$ for the cumulative distribution function of the standard normal variable, we obtain

$$x^* = \mu_U + z^*\sigma_U, \tag{9}$$

where

$$z^* = \Phi^{-1}(r) \tag{10}$$

is obtained directly from the standard normal table for the specified value of r. If $r < 1/2$, then z^* will be negative and $x^* < \mu_U$; if $r > 1/2$, then z^* will be positive and $x^* > \mu_U$.

Under the presented framework we see that lapsed funds in the public sector are equivalent to having a stockout in the newsvendor context. Lost sales are now replaced by lost funds. Meanwhile, the *off-ramp* situation equates to the newsvendor being over stocked.

4. EFFECT OF POOLING OF HOMOGENEOUS UNITS

Suppose we have two financial units denoted as unit 1 and 2. These units are assumed to be homogeneous, meaning in this context that their v_{OP}, v_{OR} and r values are the same. It would be interesting to see what happens if the *over-programming* decision for the two units is centralized at a higher level. Let U_i be the demand for over-programming dollars with mean μ_i and standard deviation σ_i, for unit i, $i = 1, 2$. Assuming these demands are independent, the combined demand,

$$U = U_1 + U_2 \tag{11}$$

is normal with mean $\mu_U = \mu_1 + \mu_2$, and standard deviation $\sigma_U = \sqrt{\sigma_1^2 + \sigma_2^2} < \sigma_1 + \sigma_2$. Following an identical marginal analysis, the pooled decision would be,

$$x^* = \mu_U + z^*\sigma_U, \tag{12}$$

where $z^* = \Phi^{-1}(r)$ has the same value as at the unit level. At the unit level, the decision would be

$$x_i^* = \mu_i + z^*\sigma_i, \quad i = 1, 2. \tag{13}$$

Thus for $z^* < 0$ ($r < \frac{1}{2}$), $\quad x^* = (\mu_1 + \mu_2) + z^*\sqrt{\sigma_1^2 + \sigma_2^2}$
$> (\mu_1 + \mu_2) + z^*(\sigma_1 + \sigma_2)$
$= x_1^* + x_2^*$. Thus, the effect of combining the two *over-programming* decisions is to

increase the total *over-programming* dollars procured. Similarly, if $z^* > 0 (r > \frac{1}{2})$, the effect would be to reduce the total procurement. The pooling effect brings the individual unit procurements closer to their mean demands.

Now consider the total value added by *over-programming*. Under centralized control, the combined demand U has a normal probability density function given by:

$$f_U(u) = \frac{1}{\sqrt{2\pi}\sigma_U} e^{-\frac{(u-\mu_U)^2}{2\sigma_U^2}}, \quad -\infty < u < \infty. \tag{14}$$

The value added under centralized control would be:

$$V_C = \begin{cases} v_{OP}x^* - v_{OR}(x^* - U), & U < x^* \\ v_{OP}x^*, & U \geq x^* \end{cases} \tag{15}$$

Thus, the expected added value is calculated as: $\mathrm{E}[\mathrm{V}_C] = \int_{-\infty}^{x^*} (v_{OP}x^* - v_{OR}(x^* - u))f_U(u)du + \int_{x^*}^{\infty} v_{OP}x^* f_U(u)du$
$= v_{OP}x^* - v_{OR} \int_{-\infty}^{x^*} (x^* - u) f_U(u)du$, which, after substituting (14) for the density function and evaluating the integral, simplifies to:

$$E[V_C] = v_{OP}x^* - v_{OR}(\frac{\sigma_U}{\sqrt{2\pi}} e^{-\frac{(z^*)^2}{2}} + (x^* - \mu_U)r). \tag{16}$$

Also note that the expected number of *off-ramp* dollars is given by:

$$E[OR_C] = \int_{-\infty}^{x^*} (x^* - u) f_U(u)du, \tag{17}$$

so that equation (4) may be rewritten as,

$$E[V_C] = v_{OP}x^* - v_{OR}E[OR_C]. \tag{18}$$

A similar analysis can be performed with units 1 and 2 operating independently (i.e., the current state of decentralized control). We obtain:

$$E[V_i] = v_{OP}x_i^* - v_{OR}(\frac{\sigma_i}{\sqrt{2\pi}} e^{-\frac{(z^*)^2}{2}} + (x_i^* - \mu_i)r), \tag{19}$$

where V_i denotes the value added by unit $i, i = 1, 2$.

We now compare the value added under the two schemes, that is, $E[V_C]$ for centralized control and $E[V_D] = E[V_1] + E[V_2]$ for the decentralized system: $\mathrm{E}[\mathrm{V}_C] - E[V_D] = v_{OP}(x^* - (x_1^* + x_2^*)) + v_{OR}((\frac{\sigma_1+\sigma_2-\sigma_U}{\sqrt{2\pi}})e^{-\frac{(z^*)^2}{2}} + (x_1^* + x_2^* - x^*)\frac{v_{OP}}{v_{OR}})$
$= v_{OR}(\frac{(\sigma_1+\sigma_2)-\sigma_U}{\sqrt{2\pi}} e^{-\frac{(z^*)^2}{2}})$
> 0, since $\sigma_1 + \sigma_2 > \sigma_U$.

Thus, the average value added by *over-programming* increases under centralized control, and hence, pooling of homogeneous units is able to improve the overall performance of the units in the long run. This result is well-known in the newsvendor context. Suppose for example that two newsvendors decide to cooperate, and

$r > 1/2$. Each newsvendor could then reduce by a few papers the over-stocking required by their individual optimal policies for this particular paper. Then, when a stockout occurs for one of the vendors, the other one can help out with any excess papers on hand.

Of course, the situation is not as simple in complex organizations as found in the public sector. Indeed, pooling of homogeneous units could be a daunting task requiring lots of planning, and ample managerial skills. However, the underlying principle remains the same. For example, lets look at the Department of National Defence (DND) of Canada. Based on limited observations, the value ratio r may be taken to be less than 1/2, and in fact appears to have a perceived value closer to 0, judging from the conservative behavior of some managers.

In this situation, lapsed funds are likely to occur in most years if managers are behaving rationally. Now suppose two units are selected to be pooled. Both would be encouraged to increase their commitments to *over-programmed* activities at the start of the year. At the end of the fiscal year, if one of them appears to be running over budget, the other unit might be able to help out if it has funds that are lapsing. In fact, this type of cooperation already exists on a limited and informal basis within DND. The goal would be to formalize and expand this process.

It would also be interesting to compare the amount of lapsed funds occurring under centralized and decentralized control. Noting that lapsed funds at year-end is given by,

$$L = max\{0, U - x^*\}, \tag{20}$$

the general formula for the expected amount of such funds is given by: E[L]=$\int_{-\infty}^{\infty} max\{0, u - x^*\} f_U(u)du$
$= \int_{x^*}^{\infty} (u - x^*) f_U(u)du$. Thus, under centralized control, we obtain

$$E[L_C] = \int_{x^*}^{\infty} (u - x^*) \frac{1}{\sqrt{2\pi}\sigma_U} e^{\frac{-(u-\mu_U)^2}{2\sigma_U^2}} du, \tag{21}$$

which after some intermediate steps simplifies to:

$$E[L_C] = \frac{\sigma_U}{\sqrt{2\pi}} e^{\frac{-(z^*)^2}{2}} + (\mu_U - x^*)(1 - r). \tag{22}$$

Under the decentralized system, total lapsed funds is given by

$$L_D = L_1 + L_2, \tag{23}$$

where the L_i denote the lapsed funds from unit $i, i = 1, 2$. It follows that

$$E[L_D] = (\frac{\sigma_1 + \sigma_2}{\sqrt{2\pi}}) e^{\frac{-(z^*)^2}{2}} + ((\mu_1 + \mu_2) - (x_1^* + x_2^*))(1 - r). \tag{24}$$

Hence, we obtain:

$$E[L_C] - E[L_D] = \frac{\sigma_U - (\sigma_1 + \sigma_2)}{\sqrt{2\pi}} e^{\frac{-(z^*)^2}{2}} + ((x_1^* + x_2^*) - x^*)(1 - r). \tag{25}$$

Note that the first term on the RHS of the preceding equation is negative, since $\sigma_U < \sigma_1 + \sigma_2$. Furthermore if $r < \frac{1}{2}$, as is the case noted above for DND, we also have $x_1^* + x_2^* < x^*$ (see(14)), and thus,

$$E[L_C] - E[L_D] < 0. \tag{26}$$

Referring to (25), we see that the reduction in lapsed funds from centralizing the *over-programming* decision is attributed to two factors:
(i) an increase in the committed amount of *over-programming* ($x^* > x_1^* + x_2^*$), and
(ii) a decrease in uncertainty of demand for *over-programming* dollars from pooling the units ($\sigma_U < \sigma_1 + \sigma_2$).

5. CONCLUSIONS

Consider the following statement by the Auditor-General of Canada [1] on the issue of lapsed funds in the Department of National Defence. *" The lack of accurate and timely information for decision makers contributed to the lapsing of more than* \$300 *million in funding that was available during the 2007-08 fiscal year but is now permanently unavailable to National Defence."* We may surmise from this statement that (i) lapsed funds are viewed in a negative way, and (ii) a better information system can solve the problem.

This paper brings into question the validity of this point of view by showing that the *"over-programming"* decision of the financial manager with respect to operating funds has the same basic ingredients of the classical newsvendor (or newsboy) problem (NVP). That is, the manager must make a decision on how many *"over-programming"* dollars to commit during the fiscal year before he/she knows with certainty how many such dollars will be available, in a similar way as the newsvendor must decide how many newspapers to order before knowing what the actual demand will be that day. Just as the goal of the NVP is to maximize expected daily profit, the rational manager will aim to maximize the expected value added by his/her *"over-programming"* decision. And just as stockouts occur from time to time in the NVP, there will be years when the manager under-spends the operating budget resulting in the loss (or lapsing) of funds that are otherwise available to his/her unit.

One could argue that comparing the financial manager to a newsvendor is a stretch, or over simplification. On the other hand, following the philosophy of Geoffrion [5] that models should provide insights, not just numbers, we argue that the NVP framework does exactly that. A better information system can perhaps alleviate the problem of lapsed funds, but it cannot eliminate the problem completely as insinuated by the Auditor-General. This is because *"over-programmed"* activities invariably take some months to complete so that the uncertainty of "demand" cannot be totally eliminated. Thus, using the NVP, we can argue that lapsed funds are in fact a good thing! That leaves us with an important, but difficult, question to answer: what is an acceptable amount of lapsed funds? The \$300 million of lapsed funds referred to above represents 2.3% of a total operating budget of \$13 billion. Is this reasonable? Or should it be 1% or $\frac{1}{2}$%? Lots of data and analysis are required before this question can be answered. However, for now our NVP analysis shows that it should not be 0%.

In summary, the simple NVP model presented here is not meant to provide numbers, but rather, to suggest strategic directions to follow through the insights provided by the model. For example, we also show in the context of the NVP that pooling homogeneous units will improve overall performance of the units, and thus, this strategy deserves further investigation.

REFERENCES

[1] Auditor General of Canada (2009). 2009 Spring Report of the Auditor General of Canada. Chapter 5, available at http://www.oag-bvg.gc.ca/internet/English/parl_oag_200905_e_32545.html

[2] Billingsly P., *Probability and Measure*, 3rd edn. John Wiley and Sons, New York, 1996.

[3] Brimberg J., and Hurley WJ., "Allocating operating funding in the public sector and the Newsvendor Problem", *Journal of the Operational Research Society*, 66 (6) (2015) 1035-1043.

[4] Brimberg J., and Hurley WJ., "Limitations on the ' better information system' solution to the problem of lapsed government operating funds", *Journal of Defense Modeling and Simulation*, 11 (4) (2014) 375-385.

[5] Geoffrion A., "The purpose of mathematical programming is insight not numbers", *Interfaces*, 7(1)(1976) 81-92.

[6] Khouja M., "The single period (News-Vendor) inventory problem: A literature review and suggestions for future research", *Omega*, 27 (5) (1999) 537-553.

[7] Petruzzi, NC., and Dada, M., "Pricing and the newsvendor problem, a review with extensions", *Operations Research*, 47 (2) (1999) 183-194.

[8] Silver, EA., Pyke, DF., and Peterson, RP., Management and Production Planning and Scheduling, 3rd edn. Wiley: New York, 1998.