How larger demand variability may lead to lower costs in the Newsboy Problem

Ad Ridder
Vrije Universiteit Amsterdam

Erwin van der Laan and Marc Salomon
Erasmus Universiteit Rotterdam

April 4, 1996

Abstract

Intuition may lead to the hypothesis that in stochastic inventory systems a higher demand variability results in larger variances and in an increase of total expected system costs. In a recent paper, Song [5] formally proved this assertion to hold for a certain class of inventory models (including the Newsboy Problem), given a particular definition of variability. Here we use stochastic dominance relations in the Newsboy Problem to characterize demand distributions for which the opposite effect may occur, i.e., higher demand variability may result in larger variances and lower costs. In addition, we provide necessary and sufficient conditions under which larger demand variances and lower costs occur simultaneously.

Keywords: Newsboy Problem, Demand Variability, Stochastic Dominance.

1 Introduction

In the Newsboy Problem a newsboy vendor decides at the beginning of the day on how many newspapers to take out to the vendor point such as to minimize total expected costs. Let Q be the number of newspapers to take out to the vendor point, D the random demand during a day, $h > 0$ the costs per newspaper in surplus at the end of the day, and $p > 0$ the penalty cost per newspaper in short at the end of the day. The cost function in the Newsboy Problem is defined as the total expected costs at the end of the day, that is

$$C(Q) = \mathbb{E}\Big(h(Q-D)^+ + p(D-Q)^+\Big), \tag{1}$$

which is minimized for

$$Q^* = \min\left\{Q = 0, 1, \ldots : \mathbb{P}(D \le Q) \ge \frac{p}{p+h}\right\}.$$

We consider two vendor points which differ only with respect to their demand processes. The demands D_1 and D_2 have equal means $\mathbb{E}D_1 = \mathbb{E}D_2$ but different variances, say $\mathrm{Var}(D_1) \le \mathrm{Var}(D_2)$. The associated optimal decisions are Q_1^* and Q_2^*. What can be said about the relation between the costs $C_1(Q_1^*)$ and $C_2(Q_2^*)$? In two recent papers [5, 6] Song has studied this question. For a specific class of demand distributions he showed that the demand with a larger variance admits higher costs. This class is identified by stochastic dominance relations, notably of low degrees. The purpose of our paper is to show that, by considering higher degrees of stochastic dominance relations, distributions may be characterized for which the opposite effect may occur.

2 Preliminaries

First, we consider the Newsboy Problem with continuous demands. In Section 5 we return to the discrete demand case. For demand D with cdf (cumulative distribution function) F we use the relations

$$\mathbb{E}(D-Q)^+ = \mathbb{E}D - Q + \mathbb{E}(Q-D)^+,$$

and

$$\mathbb{E}(Q-D)^+ = \int_0^Q F(x)\,dx \tag{2}$$

to rewrite the expected costs (1) as

$$C(Q) = p(\mathbb{E}D - Q) + (h+p)\int_0^Q F(x)\,dx. \tag{3}$$

Let us consider two systems $i = 1, 2$, where demand D_i in system i has cdf F_i and density function f_i. The cost function C_i in system i is minimized for Q_i^*. A comparison of the variability of the demands in the two systems may be carried out in many different ways, depending on the definition of variability. In [5] the following definition has been used.

Definition 1 *D_2 is more variable than D_1, denoted by $D_2 \ge_{\mathrm{var}} D_1$, if $f_2 - f_1$ changes sign exactly twice with sign sequence +,−,+.*

When D_2 is more variable than D_1 according to this definition, the mass of the distribution of D_2 is spread out more to the tails than the distribution of D_1. Hence, the variance of D_2 is larger than the variance of D_1 (we always assume equal means). Moreover, [5] proves that the expected costs (1) associated with D_2 are at least as large as those with D_1. Formally,

Lemma 1 (Song [5])

$$D_2 \geq_{\mathrm{var}} D_1 \Rightarrow C_2(Q_2^*) \geq C_1(Q_1^*).$$

In the sequel we consider alternative definitions to measure demand variability. These definitions are well accepted in the context of Utility Theory (see e.g. Fishburn and Vickson [1]). In the next section we indicate how these definitions may help us to characterize demand distributions for which the opposite effect may occur, i.e., a larger variance and lower costs.

Definition 2 *The stochastic dominance relation $\geq_n$ is defined for $n = 1, 2, \ldots$ as,*

$$D_1 \geq_n D_2 \quad \text{if and only if for all } x \geq 0 \quad H_n(x) \geq 0,$$

where

$$\begin{aligned} H_0(x) &= f_2(x) - f_1(x) \quad (x \geq 0), \\ H_n(x) &= \int_0^x H_{n-1}(t)\,dt \quad (n = 1, 2, \ldots;\ \ x \geq 0). \end{aligned}$$

The lower degrees of these dominance relations are equivalent to the stochastic orderings which are familiar in the Operations Research literature (see Chapter 1 in Stoyan [7]): $D_1 \geq_1 D_2$ is equivalent to $D_1 \geq_{\text{stochastic}} D_2$, $D_1 \geq_2 D_2$ is equivalent to $D_1 \geq_{\text{concave}} D_2$. Also, because $\mathbb{E}D_1 = \mathbb{E}D_2$, $D_1 \geq_2 D_2$ is equivalent to $D_2 \geq_{\text{convex}} D_1$. Theorem 1 in Fishburn [2] says that $(D_1 \geq_n D_2,\ \mathbb{E}D_1 = \mathbb{E}D_2,\ \mathbb{E}D_1^2 \neq \mathbb{E}D_2^2)$ implies $\mathbb{E}D_1^2 < \mathbb{E}D_2^2$. So for ease of exposition, we say that the demand D_2 is more n-variable than demand D_1, whenever D_1 stochastically dominates D_2 in degree n ($n \geq 2$).

Carrying out the proof of Lemma 1, we notice that it suffices to assume that $D_1 \geq_2 D_2$ instead of $D_2 \geq_{\mathrm{var}} D_1$. The latter is much stronger, as has been pointed out on p. 13 in [7]: $\geq_{\mathrm{var}}$ ordering implies the convex ordering $\geq_{\text{convex}}$ (and hence the $\geq_2$). To summarize,

$$D_2 \geq_{\mathrm{var}} D_1 \Rightarrow D_2 \geq_{\text{convex}} D_1 \Rightarrow D_1 \geq_2 D_2 \Rightarrow \begin{cases} \mathrm{Var}(D_1) \leq \mathrm{Var}(D_2), \\ C_1(Q_1) \leq C_2(Q_2). \end{cases}$$

3 The opposite effect

Now, let us investigate the stochastic dominance $\geq_n$ for higher degrees $n > 2$. The following theorem states a sufficient condition under which the *opposite effect* occurs, i.e., higher demand variance in combination with lower total expected costs occurs.

Theorem 1 (Sufficiency)
(i) If $D_1 \geq_n D_2$ *for some* $n > 2$, *then* $\mathrm{Var}(D_1) \leq \mathrm{Var}(D_2)$.
(ii) If the cost factors h *and* p *are such that* $H_2(Q_1^*) < 0$ *then* $C_1(Q_1^*) > C_2(Q_2^*)$.

Proof. Part *(i)* is a direct consequence of Theorem 1 in [2]. For part *(ii)* we use the definition of Q_2^* and relation (3).

$$\begin{aligned} C_2(Q_2^*) - C_1(Q_1^*) &\leq C_2(Q_1^*) - C_1(Q_1^*) \\ &= (h+p)\left(\int_0^{Q_1^*} F_2(x)\,dx - \int_0^{Q_1^*} F_1(x)\,dx\right) \\ &= (h+p)H_2(Q_1^*). \qquad \square \end{aligned}$$

Notice that the occurrence of the opposite effect depends on whether $H_2(Q_1^*) < 0$. If $H_2(x) < 0$ for some $x > 0$, then we can always construct a situation in which $H_2(Q_1^*) < 0$ by selecting a suitable set of cost parameters. Therefore, it suffices to assume that demand D_2 is more n-variable than D_1 for some $n > 2$ (but not for $n = 2$). The pairs (D_1, D_2) of demands satisfying the conditions of Theorem 1 constitute a nontrivial class, of which we provide examples in the following section.

It is also possible to give a graphical interpretation of this class. If D_2 is more n-variable than D_1 for some $n > 2$, then certainly $H_2(x) \geq 0$ on $[0, \epsilon)$ for some positive ϵ, and $H_2(\epsilon) > 0$. Requiring $H_2(\cdot)$ to become negative somewhere beyond ϵ implies at least one sign change. The same line of reasoning yields that H_1 has at least two sign changes, and, finally, that H_0 has at least three sign changes starting with the sign sequence +,–,+,–. In this way we have generalized Definition 1 where exactly two sign changes are required.

In Lemma 2 below we generalize the conditions of Theorem 1 to obtain necessary and sufficient conditions for the opposite effect to occur. For this purpose we use,

$$H_2(\infty) := \int_0^\infty H_1(x)\,dx = \mathbb{E}D_1 - \mathbb{E}D_2 = 0,$$

and consequently

$$H_3(\infty) = \frac{1}{2}\left(\mathbb{E}D_2^2 - \mathbb{E}D_1^2\right). \tag{4}$$

The latter relation is proved in the Appendix.

Lemma 2 (Necessity and sufficiency)

$$\mathrm{Var}(D_1) \leq \mathrm{Var}(D_2) \quad \textit{and} \quad C_1(Q_1^*) \geq C_2(Q_2^*)$$

$$\Leftrightarrow$$

$$H_3(\infty) \geq 0 \quad \textit{and} \quad H_2(Q_1^*) \leq \frac{p}{h+p}(Q_2^* - Q_1^*) - \int_{Q_1^*}^{Q_2^*} F_2(x)\,dx.$$

Proof.

$$\begin{aligned} C_2(Q_2^*) - C_1(Q_1^*) &= p(Q_1^* - Q_2^*) + (h+p)\left(\int_0^{Q_2^*} F_2(x)\,dx - \int_0^{Q_1^*} F_1(x)\,dx\right) \\ &= p(Q_1^* - Q_2^*) + (h+p)\left(H_2(Q_1^*) + \int_{Q_1^*}^{Q_2^*} F_2(x)\,dx\right). \end{aligned}$$

The lemma follows directly from this equality and relation (4). □

4 Examples

Traditional families of demand densities are (truncated) Normal, Lognormal, Beta, Gamma, Weibull and Uniform (Silver & Peterson [4] and Appendix B in Tijms [8]). When both densities of D_1 and D_2 are taken from one of these families, the requirements regarding the first two moments, i.e., $\mathbb{E}D_1 = \mathbb{E}D_2$ and $\mathrm{Var}(D_1) \le \mathrm{Var}(D_2)$, imply that D_2 is more 2-variable than D_1 (see the tables in Appendix 1 of [7]). For instance, suppose that $D_i \sim \mathrm{Gamma}(\lambda_i, \alpha_i)$. Then

$$\left.\begin{array}{l} \mathbb{E}D_1 = \dfrac{\alpha_1}{\lambda_1} = \dfrac{\alpha_2}{\lambda_2} = \mathbb{E}D_2 \\ \mathrm{Var}(D_1) = \dfrac{\alpha_1}{\lambda_1^2} < \dfrac{\alpha_2}{\lambda_2^2} = \mathrm{Var}(D_2) \end{array}\right\} \Rightarrow \left.\begin{array}{l} \dfrac{\alpha_1}{\lambda_1} = \dfrac{\alpha_2}{\lambda_2} \\ \alpha_1 > \alpha_2 \end{array}\right\} \Rightarrow D_2 \ge_{\mathrm{convex}} D_1 \Rightarrow D_1 \ge_2 D_2.$$

To show that the opposite effect may occur even when the demand densities belong to the same family, we consider in Example 1 the family of nonsymmetric triangular densities.

Example 1
A nonsymmetric triangular density f is characterized by three parameters $a < b < c$,

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)} & \text{on } [a,b], \\ \dfrac{2(c-x)}{(c-a)(c-b)} & \text{on } [b,c], \\ 0 & \text{else.} \end{cases}$$

To fulfill the conditions of Theorem 1 and/or Lemma 2, the parameters a_i, b_i, c_i of demand density f_i satisfy a set of (in)equalities which follow from the expressions of the mean, the variance, and the H_n functions for $n = 0, 1, 2, 3$. From the feasible solutions of this set we present below two examples. The first one relates to Theorem 1: demand D_2 is more 3-variable than D_1 (and not more 2-variable). The second example relates to Lemma 2 where the opposite effect occurs when the conditions of Theorem 1 do not apply.

(a) Let the triangular densities be characterized by

$$a_1 = 1, b_1 = 2, c_1 = \frac{11}{2}; \qquad a_2 = 0, b_2 = 4, c_2 = \frac{9}{2},$$

and let the cost factors be $h = 1$ and $p = 6$. The numerical results are presented in Table 1.

i	Q_i^*	$H_2(Q_i^*)$	$\mathbb{E}D_i$	$\mathrm{Var}(D_i)$	$C_i(Q_i^*)$	$H_3(\infty)$
1	4	-0.0529	17/6	0.9306	1.6667	1/24
2	3.9279	-0.0546	17/6	1.0139	1.2883	

Table 1. Results for Example 1(a)

The graphs of $H_0, \ldots, H_3$ are easily obtained. They show three sign changes of H_0, and $H_3 \geq 0$, implying $D_1 \geq_3 D_2$. Clearly, also the conditions of Lemma 2 must be satisfied. Indeed, calculus yields

$$H_2(Q_1^*) = -0.0529 < 0.0011 = \frac{p}{h+p}(Q_2^* - Q_1^*) - \int_{Q_1^*}^{Q_2^*} F_2(x)\,dx.$$

(b) Let

$$a_1 = 0, b_1 = 4, c_1 = 6; \qquad a_2 = 1, b_2 = 2, c_2 = 7,$$

and the cost factors $h = 5$ and $p = 1$.

i	Q_i^*	$H_2(Q_i^*)$	$\mathbb{E}D_i$	$\mathrm{Var}(D_i)$	$C_i(Q_i^*)$	$H_3(\infty)$
1	2	$-1/18$	10/3	28/18	2	1/12
2	2	$-1/18$	10/3	31/18	5/3	

Table 2. Results for Example 1(b)

In this example $H_0 < 0$ on $(0, 1)$, hence all $H_n < 0$ on $(0, 1)$. We cannot expect the existence of an n such that $H_n \leq 0$ everywhere, since that would imply $D_2 \geq_n D_1$ and $\mathrm{Var}(D_2) \leq \mathrm{Var}(D_1)$, which is not so. Therefore, it is never true that D_1 is more n-variable than D_2, or that D_2 is more n-variable than D_1. On the other hand, straightforward calculus shows that the conditions of Lemma 2 are fulfilled:

$$H_3(\infty) = \frac{1}{12} > 0, \quad H_2(Q_1^*) = -\frac{1}{18} < 0 = \frac{p}{h+p}(Q_2^* - Q_1^*) - \int_{Q_1^*}^{Q_2^*} F_2(x)\,dx.$$

Next we consider an example in which the demand density functions belong to different families.

Example 2
Let D_1 have a Lognormal ($\mu = -0.1, \sigma^2 = 0.2$) density and D_2 a Gamma ($\lambda = 4$, $\alpha = 4$) density, i.e.,

$$f_1(x) = \frac{1}{x\sqrt{0.4\pi}} e^{-(\log x + 0.1)^2/0.4}, \qquad f_2(x) = \frac{4}{6}(4x)^3 e^{-4x} \quad (x > 0).$$

Furthermore, let $h = 1$, and $p = 24$. Table 3 summarizes the numerical results.

i	Q_i^*	$H_2(Q_i^*)$	$\mathbb{E}D_i$	$\mathrm{Var}(D_i)$	$C_i(Q_i^*)$	$H_3(\infty)$
1	1.9797	-0.0013	1	0.2214	1.4052	0.0143
2	2.0214	-0.0014	1	0.25	1.3712	

Table 3. Results for Example 2

The graph of H_0 shows three sign changes in sequence +,–,+,–. The integrals H_1, H_2, H_3 are determined numerically and show that $H_3 \geq 0$, implying that D_2 is more 3-variable than D_1. Furthermore, calculus yields

$$H_2(Q_1^*) = -0.001442 < 0.000097 = \frac{p}{h+p}(Q_2^* - Q_1^*) - \int_{Q_1^*}^{Q_2^*} F_2(x)\,dx,$$

in accordance with Lemma 2.

5 The finite discrete model

In the discrete version of the Newsboy Problem we assume finite discrete demands, say

$$D_i \in I = \{0, 1, \ldots, M\},$$

with densities f_i, i.e., $f_i(k) = \mathbb{P}(D_i = k)$. The optimal decisions Q_1^* and Q_2^* are integer valued. The integral in the right hand side of (3) is replaced by

$$\mathbb{E}(Q - D)^+ = \sum_{k=0}^{Q-1} \mathbb{P}(D \leq k).$$

The stochastic dominance version of Definition 2 has been identified in Fishburn and Lavalle [3].

Definition 3 *The stochastic dominance relation $\geq_n^I$ is defined for $n = 1, 2, \ldots, M$ as,*

$$D_1 \geq_n^I D_2 \quad \textit{if and only if}$$

$$\begin{cases} h_n \geq 0, & n = 1, 2 \\ h_n \geq 0 \textit{ and } h_k(M - k + 1) \geq 0, \quad k = 2, \ldots, n-1, & n = 3, \ldots, M, \end{cases}$$

where

$$\begin{aligned} h_0(k) &= f_2(k) - f_1(k), \\ h_n(k) &= \sum_{\ell=0}^{k} h_{n-1}(\ell) \quad (n = 1, 2, \ldots, M). \end{aligned}$$

Applying Corollary 2 of [3] we adapt Theorem 1 as follows.

Theorem 2
(i) If $D_1 \geq_n^I D_2$ *for some* $n = 3, 4, \ldots, M$*, then* $\mathrm{Var}(D_1) \leq \mathrm{Var}(D_2)$.
(ii) If the costs factors h *and* p *are such that* $h_2(Q_1^* - 1) < 0$ *then* $C_1(Q_1^*) > C_2(Q_2^*)$.

Example 3
Let $I = \{1, 2, 3, 4\}$ and let the demand densities be

	1	2	3	4
f_1	1/15	6/15	1/15	7/15
f_2	3/15	1/15	5/15	6/15

The demands satisfy $h_3(1) = 2/15,\ h_3(2) = 1/15,\ h_2(3) = 0$, which implies that $D_1 \geq_3^I D_2$. For cost factors $h = 1$ and $p = 1$ the numerical results are listed in Table 4.

i	Q_i^*	$h_2(Q_i^* - 1)$	$\mathbb{E}D_i$	$\mathrm{Var}(D_i)$	$C_i(Q_i^*)$
1	3	$-1/15$	44/15	254/225	1
2	3	$-1/15$	44/15	284/225	13/15

Table 4. Results for Example 3

6 Conclusion

The conclusion of this paper is that a reduction of the demand uncertainty in stochastic production and inventory models is not always economically favourable. Whether uncertainty reduction indeed results in cost reductions depends on many factors such as the definition of uncertainty, the structure of the demand distributions, and the ratio between the shortness and surplus costs. We have formally proved this for the classical Newsboy Problem. However, the same conclusion holds for dynamic inventory models controlled by a base stock policy, where D in the cost function (1) stands for the lead time demand and Q for the base stock level.

References

[1] P.C. Fishburn and R.G. Vickson (1978). Theoretical foundations of stochastic dominance. In *Stochastic dominance*, eds. G.A. Whitmore and M.C. Findlay, Heath, Lexington, Mass., p. 39 – 114.

[2] P.C. Fishburn (1980). Stochastic dominance and moments of distributions. *Mathematics of Operations Research* 5, p. 94 – 100.

[3] P.C. Fishburn and I.H. Lavalle (1995). Stochastic dominance on unidimensional grids. *Mathematics of Operations Research* 20, p. 513 – 525.

[4] E.A. Silver & R. Peterson (1985). *Decision systems for inventory management and production planning.* Wiley, New York.

[5] J-S. Song (1994). The effect of leadtime uncertainty in a simple stochastic inventory model. *Management Science* 40, p. 603 – 613.

[6] J-S. Song (1994). Understanding the lead-time effects in stochastic inventory systems with discounted costs. *Operations Research Letters* 15, p. 85 – 93.

[7] D. Stoyan (1983). *Comparison methods for queues and other stochastic models.* Wiley, New York.

[8] H.C. Tijms (1994). *Stochastic models. An algorithmic approach.* Wiley, new York.

Appendix

The calculation of $H_2(\infty)$ is easy. Let $\overline{F}_i$ be the tail distribution function of demand D_i, $\overline{F}_i(x) = 1 - F_i(x)$. Then

$$H_2(\infty) = \int_0^\infty (F_2(x) - F_1(x))\, dx = \int_0^\infty \left(\overline{F}_1(x) - \overline{F}_2(x)\right)\, dx = \mathbb{E}D_1 - \mathbb{E}D_2.$$

For the calculation of $H_3(\infty)$ we make use of the so-called excess equilibrium distribution functions in Renewal Theory,

$$G_i(x) := \frac{1}{\mathbb{E}D_i} \int_0^x \overline{F}_i(y)\, dy \quad (x \geq 0).$$

Let X_i be a nonnegative random variable with distribution function G_i. The expectation of X_i is well known to be

$$\mathbb{E}X_i = \int_0^\infty \overline{G}_i(x)\, dx = \frac{\mathbb{E}D_i^2}{2\mathbb{E}D_i}.$$

Hence, since $\mathbb{E}D_1 = \mathbb{E}D_2$,

$$
\begin{aligned}
H_3(\infty) &= \int_0^\infty H_2(x)\,dx \\
&= \int_0^\infty \int_0^x (F_2(y) - F_1(y))\,dy\,dx \\
&= \int_0^\infty \int_0^x \left(\overline{F}_1(y) - \overline{F}_2(y)\right)\,dy\,dx \\
&= \int_0^\infty \Big((\mathbb{E}D_1)G_1(x) - (\mathbb{E}D_2)G_2(x)\Big)\,dx \\
&= \int_0^\infty \Big((\mathbb{E}D_1)(1-\overline{G}_1(x)) - (\mathbb{E}D_2)(1-\overline{G}_2(x))\Big)\,dx \\
&= \int_0^\infty (\mathbb{E}D_2)\overline{G}_2(x)\,dx - \int_0^\infty (\mathbb{E}D_1)\overline{G}_1(x)\,dx \\
&= \frac{1}{2}\left(\mathbb{E}D_2^2 - \mathbb{E}D_1^2\right).
\end{aligned}
$$