Analysis of Three Crop Risks: Demand, Yield and Harvest

Stuart J. Allen

Situation:

1. We must decide how many acres, **Q**, to plant for a single crop when faced with uncertain demand, **D**, and uncertain yield, **Y**.
2. Upon crop maturation, we must decide the rate at which to harvest, **R** , when faced with an uncertain length of time available for harvesting, **L**.

Some Definitions:

p = market price of the crop in \$/unit

c_P = variable cost of crop from planting and maintenance until maturity in \$/unit

c_R = variable cost of harvesting in \$/unit

v = salvage value of excess crop in \$/unit

It is also clear from earlier definitions that

QY = crop size, units

RL = total amount of crop harvested, units

Note that if RL > QY, the entire crop is harvested but that an opportunity cost is incurred in the form of an excessive harvest rate. The amount harvested cannot exceed QY, however. On the other hand if RL < QY, some of the crop will be left unharvested as the season ends due to weather or over-maturation. Additional opportunity costs in both of these cases will depend on the demand, D, and we will examine these in detail next.

Opportunity Costs:

Please refer to Figure 1, below, which illustrates the five important regions within the three-dimensional coordinate frame defined by the 3 random variables L, Y, D. We will look at each region separately. For clarity, Figure 1 shows only a finite volume in the positive octant. All points lying above the blue plane (regions 1 and 5) have D > QY and all below (regions 2,3 and 4) have D < QY. All points behind the green plane (i.e. in the negative L-direction, regions 1, 2 and 3) have RL < QY and in front of that plane (regions 4 and 5), RL > QY. Finally, points above the

red plane (regions 1 and 2), have D > RL and below the red plane (region 3), D < RL.

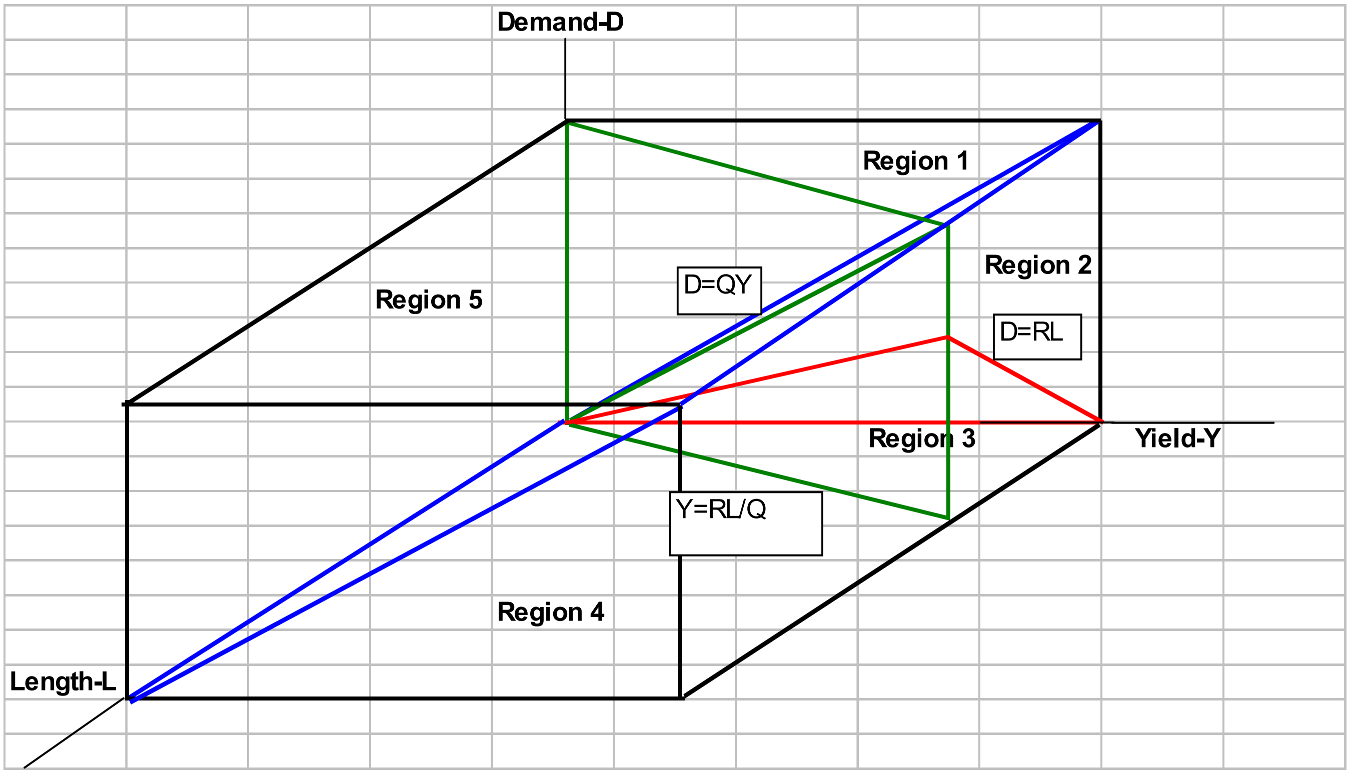

Figure 1. The Five Regions Governing Opportunity Costs

Note that the red plane does not extend through the green plane into the regions for which RL > QY since

$$\text{Crop Size} = \begin{cases} \text{QY if RL} \geq \text{QY} \\ \text{RL if RL} < \text{QY} \end{cases}$$

Furthermore, the red plane must lie everywhere below the blue plane in the region where

RL < QY. We are now ready to look at the details of the opportunity costs.

Region 1 : RL < QY ; D > QY > RL

Here we fail to gather the entire crop, so that the actual harvest size is RL. Demand exceeds both the available crop and the amount retrieved. Costs are associated with lost sales and we define these in two parts as :

$$\textbf{Lost sales} = (p-c_R) (QY - RL) + (p - c_R - c_P) (D - QY) .$$

The first term represents the sales lost in failing to retrieve the entire crop with market price reduced by the cost of the additional harvesting effort required. However, we do not use the variable crop costs as a penalty in this case since this portion of the crop already exists. The second term reduces the market value by the effort required to both

plant and harvest to meet the demand. This can be rearranged into the form

$$\textbf{Lost sales} = (p - c_R) (D - RL) + c_P (QY - D) \qquad (1)$$

which will be useful when we look at cost expectations later.

Region 2 : RL < QY ; QY > D > RL

In this region, opportunity costs are of two types: lost sales and excess crop size which incurs unnecessary costs of planting and crop maintenance through maturity. We refer to this latter by the simpler term " over-planting ".

$$\textbf{Lost sales} = (p - c_R) (D - RL) \qquad (2a)$$

$$\textbf{Over-plant} = c_P (QY - D) \qquad (2b)$$

These costs reflect our earlier discussion of Region 1 costs.

Region 3 : RL < QY ; QY > RL > D

Here the harvest is smaller than the crop and greater than demand. So we have an excess of supply (hence excess harvesting costs) which may be mitigated by salvage. In addition, we again have an over-planted condition with attendant opportunity costs.

$$\textbf{Excess supply} = (c_R - v) (RL - D) \qquad (3a)$$

$$\textbf{Over-plant} = c_P (QY - D) \qquad (3b)$$

The term $c_P (QY - D)$ occurs throughout all three regions for the case **RL < QY** and the term **(p – cR) (D – RL)** is common to regions 1 and 2. These observations can be used to simplify our formulations of cost expectations later.

Region 4 : RL > QY ; D < QY

The entire crop has been harvested so the available supply is **QY**. However, in region 4 , supply exceeds demand. Two different opportunity costs must be accounted for here: an excess harvest rate and an excess of supply.

$$\textbf{Excess supply} = (c_R + c_P - v) (QY - D) \qquad (4a)$$

$$\textbf{Excess harvest rate} = c_R (RL - QY) \qquad (4b)$$

Region 5 : RL > QY ; D > QY

The opportunity costs in this region are due to lost sales and an excessive harvest rate.

$$\textbf{Lost sales} = (p - c_R - c_P) (D - QY) \qquad (5a)$$

$$\textbf{Excess harvest rate} = c_R (RL - QY) \qquad (5b)$$

The excess harvest rate penalty is common to regions 4 and 5.

Assumptions Concerning the Random Variables L, Y, D :

We assume these three random variables are statistically independent, and normally distributed:

$$\mathbf{L \sim N(\mu_L , \sigma_L^2)\ ; \quad Y \sim N(\mu_Y , \sigma_Y^2)\ ; \quad D \sim N(\mu_D , \sigma_D^2)}$$

Dimensionless Decision Variables

It is convenient to work with dimensionless decision variables for acreage planted, Q, and harvest rate, R. First define "risk-free" values for each of these variables:

Risk-free acreage = $Q_0 = \mu_D/\mu_Y$; Risk-free harvest rate = $R_0 = Q_0\mu_Y/\mu_L$

These represent reasonable values for Q and R if there were no uncertainty in demand, yield and length of harvest season. We can then define dimensionless decision variables as:

$$\hat{Q} = Q / Q_0 \quad ; \quad \hat{R} = R / R_0$$

These definitions absorb many of the statistical parameters and have the further advantage that any search for optimal values of $\hat{Q}$ and $\hat{R}$ can begin in the neighbor hood of 1.0. It is also useful to define three coefficients of variation

$$k_L = \sigma_L/\mu_L \ ; \ k_Y = \sigma_Y/\mu_Y \ ; \ k_D = \sigma_D/\mu_D \ .$$

Expectations

In forming the expected values of the various quantities defined above, it will be useful to employ the notation E(.)i where the subscript(s) i will refer to the individual region(s).

E(Q – Y)$_{123}$:

This expected value holds for all three regions 1, 2, and 3 hence the triple subscript. The random variables have the following ranges

$$RL/Q \le Y < \infty \ ; \ -\infty < D < \infty \ ; \ -\infty < L < \infty \ .$$

Then

$$E(Q-Y)_{123} = \int_{L=-\infty}^{\infty} \int_{D=-\infty}^{\infty} \left\{ \int_{Y=RL/Q}^{\infty} (QY-D) f(Y) \right\} f(D) f(L) dD dL$$

After some algebra we obtain

$$E(Q-Y)_{123} = \mu_D \int_{z_L=-\infty}^{\infty} \left\{ (\hat{Q}-1)[1-\Phi(K)] + \hat{Q}\, k_Y\, \varphi(K) \right\} \varphi(z_L) dz_L$$

where $\Phi(K)$ and $\phi(K)$ are the standard normal cumulative and density functions respectively and are evaluated at K, defined by:

$$K = [(\hat{R}/\hat{Q})(1 + k_L z_L) - 1]/k_Y$$

E(D – RL)$_{12}$:

Here the random variables are in the ranges

$$RL/Q \le Y < \infty \ ; \ RL \le D < \infty \ ; \ -\infty < L < \infty$$

So

$$E(D-RL)_{12} = \int_{L=-\infty}^{\infty} \int_{D=RL}^{\infty} \left\{ \int_{Y=RL/Q}^{\infty} (D-RL) f(Y) dY \right\} f(D) f(L) dD dL$$

Again we can carry out integration with respect to Y and D giving

$$E(D-RL)_{12} = \mu_D k_D \int_{z_L=-\infty}^{\infty} \left[1-\Phi(K)\right] \left\{ -M\left[1-\Phi(M)\right] + \varphi(M) \right\} \varphi(z_L) dz_L$$

where

$$M = \left[\hat{R}\left(1 + k_L z_L\right) - 1\right] / k_D$$

E(RL – D)$_3$:

The region of interest here is:

$$RL/Q \le Y < \infty \;\; ; \;\; -\infty < D \le RL \;\; ; \;\; -\infty < L < \infty$$

We next form the expectation

$$E(RL - D)_3 = \int_{L=-\infty}^{\infty} \int_{D=-\infty}^{RL} \left\{ \int_{Y=RL/Q}^{\infty} (RL - D) f(Y) dY \right\} f(D) f(L) dD dL$$

with final result

$$E(RL - D)_3 = \mu_D k_D \int_{z_L=-\infty}^{\infty} [1 - \Phi(K)][M\Phi(M) + \varphi(M)]\varphi(z_L) dz_L$$

E(RL – QY)$_{45}$:

For this region:

$$-\infty < Y \le RL/Q \;\; ; \;\; -\infty < D < \infty \;\; ; \;\; -\infty < L < \infty$$

$$E(RL - QY)_{45} = \int_{L=-\infty}^{\infty} \int_{D=-\infty}^{\infty} \left\{ \int_{Y=-\infty}^{RL/Q} (RL - QY) f(Y) dY \right\} f(D) f(L) dD dL$$

This becomes, after integration on Y and D:

$$E(RL - QY)_{45} = \mu_D k_Y \hat{Q} \int_{z_L=-\infty}^{\infty} [K\Phi(K) + \varphi(K)]\varphi(z_L) dz_L$$

E(QY – D)$_4$:

The regions for this expectation are:

$$-\infty < D \le QY \;\; ; \;\; -\infty < Y \le RL/Q \;\; ; \;\; -\infty < L < \infty$$

$$E(QY - D)_4 = \int_{L=-\infty}^{\infty} \int_{Y=-\infty}^{RL/Q} \left\{ \int_{D=-\infty}^{QY} (QY - D) f(D) dD \right\} f(Y) f(L) dY dL$$

Integration over D and (subsequently) Y yields

$$E(QY - D)_4 = \mu_D k_D \int_{z_{L=-\infty}}^{K} [I_1(z_L) + I_2(z_L)]\varphi(z_L)dz_L$$

where

$$I_1(z_L) = \int_{z_Y=-\infty}^{K} \varphi(N)\varphi(z_Y)dz_Y \quad \text{and} \quad I_2(z_L) = \int_{z_{Y=-\infty}}^{K} N\Phi(N)\varphi(z_Y)dz_Y$$

while

$$N = [\hat{Q}(1 + k_Y z_Y) - 1]/k_D = az_Y + b \ ; \ a = \hat{Q}k_Y/k_D \ ; \ b = (\hat{Q} - 1)/k_D$$

These two integrals on z_Y present some complications for spreadsheet implementation of a numerical integration process and we will address these later after we obtain the last expected value.

E(D – QY)$_5$:

Now the region is defined by:

$$QY \le D < \infty \ ; \ -\infty < Y \le RL/Q \ ; \ -\infty < L < \infty$$

$$E(D - QY)_5 = \int_{L=-\infty}^{\infty} \int_{Y=-\infty}^{RL/Q} \left\{ \int_{D=QY}^{\infty} (D - QY)f(D)dD \right\} f(Y)f(L)dYdL$$

Integration on D and then Y gives

$$E(D - QY)_5 = \mu_D k_D \int_{zL=-\infty}^{\infty} [-\Phi(K)(\hat{Q} - 1)/k_D + \varphi(K)k_Y/k_D + I_1(z_L) + I_2(z_L)]\varphi(z_L)dz_L$$

where the $I_i(z_L)$ are as defined previously. Next we will turn our attention to a more suitable form for these two integrals.

Reduction of the $I_1(z_L)$ Integral

First, using the earlier definition of N, we form

$$N^2 + z_Y^{\ 2} = \left[1 + \left(\frac{\hat{Q}k_Y}{k_D}\right)^2\right] z_Y^{\ 2} + 2\frac{k_Y}{k_D^{\ 2}}\hat{Q}(\hat{Q}-1)z_Y + \frac{1}{k_D^{\ 2}}(\hat{Q}-1)^2$$

We can rewrite $(N^2 + z_Y^{\ 2})$ as

$$N^2 + z_Y^{\ 2} = \left[\sqrt{1+a^2}\, z_Y + ab/\sqrt{1+a^2}\right]^2 + \left[b/\sqrt{1+a^2}\right]^2$$

Now set

$$w = \sqrt{1+a^2}\, z_Y + ab/\sqrt{1+a^2} \quad ; \quad \hat{K} = \sqrt{1+a^2}\, K + ab/\sqrt{1+a^2}$$

so that

$$I_1(z_L) = \frac{1}{\sqrt{1+a^2}\sqrt{2\pi}} \exp[-(b/\sqrt{1+a^2})^2/2] \int_{w=-\infty}^{\hat{K}} \frac{1}{\sqrt{2\pi}} \exp(-w^2/2)dw$$

The integral has been reduced to the cumulative at $\hat{K}$ which depends on z_L (see the definition of K earlier). Finally

$$I_1(z_L) = \varphi(b/\sqrt{1+a^2})\,\Phi(\hat{K})/\sqrt{1+a^2}$$

Reduction of the $I_2(z_L)$ Integral

This integral separates naturally into two parts:

$$I_2(z_L) = aI_3(z_L) + bI_4(z_L)$$

where

$$I_3(z_L) = \int_{-\infty}^{K} z\Phi(az+b)\varphi(z)dz \qquad .$$

$$I_4(z_L) = \int_{-\infty}^{K} \Phi(az+b)\varphi(z)dz$$

We deal first with $I_3(z_L)$ using integration by parts with $u = \Phi(az+b)$ and

$v = -\phi(z)$, then

$$I_3(z_L) = -\varphi(K)\Phi(aK+b) + aI_2(z_L)$$

where $I_2(z_L)$ has been simplified earlier. Thusfar, $I_4(z_L)$ has eluded simplification.