Chapter 7: Risk and Market Outcomes

I. Competition and Profits

In the choice settings modeled in Part I, the decision makers have been assumed to have clear overarching goals (utility, net benefits, or profit) and complete information about their possibilities. In such settings, no mistakes are ever made by rational decisionmakers, thus consumers choices always truly maximize their utility and firms always maximize their profits. In markets where a large number of firms (and potential firms) compete for the purchases of consumers, rates of return across markets all tend to converge to the same rate. Any market with above average rates of return attracts entry by both new firms and firms selling in other markets where returns are lower. In markets where firms earn below average rates of return, firms tend to exit such markets and, to the extent possible, shift their resources to ones earning average or above average returns. Thus, in equilibrium rates of return equilibrate across markets. This is sometimes referred to as a "zero profit" equilibrium because firms (and firm owners) are just earning their opportunity cost rate of return on their investments (although rates of return are positive).

A similar process is imagined taking place among purchasers and sellers of goods and services. They purchase their goods from suppliers with the lowest price. This causes firms with above average prices to lose sales (in the limit all sales) and encourages them to lower prices to the lowest that is sustainable (e.g., they produce at the lowest point on their long run average cost curve). Such pricing pressures tend to cause all prices to converge to a single price equal to each firm's lowest long run average cost. In this manner all the firms in the market have to be efficient—indeed equally efficient—and thus all are efficiently sized. This is also the type of setting Marshall's models of long run supply and long run price determination were developed to explain. The only limitations to this theory are the assumptions that a relatively large number of firms can be supported (enough to make the price-taking assumption plausible) and that both firms and consumers have sufficient knowledge that they do not make mistakes. General equilibrium theory, some decades later, would demonstrate that prices can in such circumstances simultaneously clear all markets, including input markets and markets for loans.

Divergence from this environment of "perfect competition," can thus take place for two reasons. First, there may be sufficient economies of scale (possibly caused by relatively fixed costs) that only a small number of firms can be supported in equilibrium. Second, consumers and firms may not know enough to make mistake-free decisions. Significant ignorance may exist about technological possibilities, the prices available at rival sellers, and the profits realized by firms selling in different markets. Third, there may be various entry barriers that limit entry and exit possibilities both for new and existing firms. The first type of phenomena helps explain variation in market structure or industrial organization. The second type of phenomena explains why variations among firms and consumers tend to exist, even if employees, owners, and consumers are essentially similar beings. The third suggests that government policies may be important determinants of the extent and efficiency of markets in a given country.

This chapter deals with the second of these phenomena. The third is taken up in part three of the book. The first was addressed to some extent in the second part of chapter 3, where firms were price makers rather than price takers. Industrial organization is a large, specialized field, and complete coverage of that field is beyond the intent of this book.

Statistical theory provides us with a vocabulary to describe choice settings in which decision makers are less than entirely certain of the consequences of their decisions. In some of these settings, one may be able to describe the consequences with well-known probability functions. In that setting at least some of the relevant consequences are generated by stochastic processes or other processes that have consequences that can be described (with reasonable accuracy) as if they were generated by such phenomena. (The "random" number generators of computers are examples of such phenomena.)

Although persons can be said to be less than perfectly informed in such circumstances, because they do not know the precise consequences of their choices or all of the relevant characteristics of their choice settings, there is another sense in which they can be said to know as much as can be known. If the process is truly stochastic, then all that can be known about it is its probability function or probability density function.

Nonetheless, it also can be said that the characteristics of the relevant probability function or probability density function determine just how certain one can be about one's understanding of one's choice setting and the consequences of one's choices. The broader the variance in the probability distribution that describes the consequences of one's choices the less certain one is about the consequences of those choices.

Frank Knight added another category to this spectrum of choice settings, namely settings in which one does not know the distribution of the probability distribution generating consequences, either because of ignorance or because the phenomenon itself is so idiosyncratic that it cannot be described with a probability distribution. The latter settings he terms "uncertain" and the former settings "risky." Risky choices can be completely modelled, whereas uncertain choices cannot—or to the degree that they can be modelled they are less sharply modelled. Knight argues that all profits (and losses) realized in competitive markets arise because of uncertainty rather than risk for reasons that we'll discuss in this chapter.

II. Expected Values and Expected Utility

A modest extension of the rational choice model can be used to characterize decisions in choice settings in which the probability function associated with some phenomena is both generally well-understood and well-known. Namely, it is assumed that consumers maximize "expected" utility and firms maximize "expected" profits. In many areas of choice, the benefits and costs of particular choices (or policies) are at least partly the consequence of chance.

The most common models of decision making in settings of uncertainty are the expected utility and expected net benefit maximizing models. The notion of "**expected value**," itself, is an idea taken from statistics and means the average result that would be expected from a series of "draws" from a stable random process of some kind.

DEF: The **mathematical expected value** of a set of possible outcomes, 1, 2, ... N with values $V_1, V_2, \dots V_N$ and probabilities of occurrence $P_1, P_2, \dots P_N$ is:

$$E(V) = \sum_{i=1}^{N} P_i V_i$$

Every **probability function** assigns probabilities to discrete events (here events 1, 2, … N) such that the sum of the probabilities is 1.0. (The probability that something will actually happen is 1, is completely certain.) Every probability distribution has the property: $\Sigma Pi = 1$ with $P_i \geq 0$ for all i . Every possibility is assigned a probability. All possibilities, i, have positive probabilities of occurrence $1 \geq Pi > 0$. All impossibilities, j, have a zero probability of occurring and so $Pj = 0$. The **mathematical expected value is the sum of the values of those possibilities (here $V_1, V_2 \dots$**

V_N) times their particular probabilities of occurrence (here P_1, P_2, … P_N). It represents the long-term *average* value of the distribution of values.

Expected utility, thus, is a special case of expected values, namely cases in which the "value" is measured in terms of utility (as utils). The *expected* utility associated with a probabilistic setting is calculated is thus in a similar manner:

$$E(V) = \sum_{i=1}^{N} P_i U(v_i)$$

where the N "value possibilities" are now measured in utility terms associated with the affected individual. To use this formula for expected utility calculations, one has to assume that the outcomes of the "uncertain" events are finite, can be counted, can be listed, and probabilities assigned to them.

This is not an unreasonable assumption in many circumstances and is a reasonable first approximation of many others. The probabilities assigned may be the result of careful empirical work (frequentist) or (Bayesian) intuitions about the likelihood of particular events that are updated as more evidence is gathered. This latter case is one way to use probabilistic choice models to think about choice settings that Knight would regard to be uncertain. This possibility is taken up in chapter 7.

Most economists and most economic models are quite willing to assume that all the possible outcomes are known, that probabilities can be assigned to them, and that the possibilities are countable or be modelled as countable. However, there are statistical tools for dealing with probabilistic phenomena whose outcomes are not countable, as taken up in the next section of this chapter. In most policy areas, however, the probabilities are themselves estimates that are updated as research, policies, or persuasive campaigns take place.

Illustration of the Difference Between Expected Values and Expected Utility

To illustrate the difference between expected values and expected utility consider, the expected roll of a die (a single dice). Suppose that a single die is to be rolled. The face that turns up on top is a random event. Suppose that you will be paid a dollar amount equal to the number on the face that winds up on top. Since the probability of a particular face winding up on top is 1/6 and the value of the outcomes are 1, 2, 3, 4, 5, 6, arithmetic implies that the expected value of this game in money terms is \$3.50 = (1)(1/6) + (2)(3.5) + (3)(1/6) + ...(6)(1/6). If you played the game dozens of times, your average payoff per roll would be approximately\$3.50.

Note that the expected value of a single roll of a die is 3.5, a number that actually is impossible, rather than "expected" in the usual sense in ordinary English. This is not always the case, but this example illustrates that the meaning of "expected value" is a technical one: namely the long-term average result, rather than the result that you would most commonly observe. There are many probability distributions in which the average value is also the mode, as with the normal distribution, in which case the expected value the most likely value to be observed,

Next, we'll repeat the exercise for a concrete utility function, namely U =$V^{.5}$, where V is the winnings from a particular roll of the die. In this case

$$E(U(V)) = \sum_{i=1}^{N} (1/6)\,(i)^{.5}$$

Or $U^e = (0.1667) + (0.2357) + (0.2887) + (0.3333) + (0.3726) + (4082) = 1.8053$

Note also that if the intermediate cases were for some reason impossible—or simply ignored as far as prizes are concerned, there would be just two possibilities, each with a probability of .5. In that case the expected utility is:

$$U^e = .5\,(1)^{.5} + .5(6)^{.5} = (.5) + (1.2247) = 1.7247 < 1.8053$$

Expected utility falls because the stochastic event becomes "riskier" because only the extreme possibilities are possible in the second case. Note that the expected value of the "gamble" or choice circumstances is 3.5 in each case. Al, as we shall see later in the chapter, is quite risk averse.

Utility functions that can be used to calculate expected utility values that properly rank alternative outcomes (according to expected utility) are called **Von-Neumann Morgenstern utility functions**. Von-Neuman Morgenstern utility functions are all complete, transitive, continuous, and exhibit monotonicity. In addition, they have the property of what is sometimes called substitutability which might also be considered internal consistency with respect to stochastic circumstances. If one is indifferent between x and y then one is also indifferent between px and (1-p)z and py and (1-p)z, where p is the probability of event x. And if z is regarded to be better than x, than pu(x) + (1-p)u(y) < pu(x) + (1-p)u(z).

Experiments have been undertaken to use various gambles to create Von Neuman – Morganstern Utility functions—which, as it turns out, do not perfectly explain individual behavior under uncertainty in laboratories, but do so reasonably well. Von-Neuman Morgenstern utility functions for particular individuals are also "unique" up to a linear transformation (and considered by some to be a form of cardinal utility), because one can do arithmetic with them.

Expected Utility with Continuous Probability Functions

The above is developed for cases in which the stochastic factor(s) is (are) take only discrete values that are countable, such as the value generated when rolling dice or yes/no types of outcomes, such as one's house burns down or does not or being fired or not, or being fined or not for violating a law or regulation. Expected values for continuous cases—choice settings where the domain of possibilities is compact—can be represented with integrals.

Suppose that Al has a two-good strictly concave utility function, U=u(A,B) where the prices of goods A and B are P_A and P_B respectively. Al has W dollars to spend in the period of interest. The quality of good A is not known at the point of purchase, whereas that of good B is known with certainty. Given f(q), the frequency distribution of the quality of some good that an individual may purchase, the expected utility for Q units of good A can be written as an integral of the following sort. (Note that the substitution "trick" has been used to characterize good B as a function of purchases of good A.)

$$\mathrm{U}^e = \int_L^H \mathrm{U}(\mathrm{Q}(\mathrm{q}), (W - P_A Q)/P_B)\mathrm{f}(\mathrm{q})\ \ d\mathrm{q}$$

The integral written above is the expected (or average) marginal utility associated with output Q. The quantity that maximizes expected utility can be found by differentiating expected utility with respect to Q and setting the result equal to zero. The first order condition in this case takes the form:

$$\mathrm{U}_Q^e = \int_L^H [U_A f(q) - U_B(\frac{P_A}{P_B})] d\mathrm{q} = 0 \ \text{ at } Q^*$$

where A and B subscripts indicate derivatives with respect to the variables subscripted.

Notice that the partial derivatives are results obtained by differentiating the integrand. The integral domains are carried forward and the function being integrated (the integrand) is replaced with its relevant first derivatives. And also notice that what one obtains in this case are terms for the expected marginal benefit (the integral of the first term, in terms of utils) and for the expected cost of units of A (the integral of the second term, again in terms of utils).[1]

The implicit function theorem implies that the Al's demand for good A, here Q*, can be written as $Q^* = g(P_A, P_B, W, H, L)$. The effect of the quality uncertainty affects the shape of this function, but, as written, the demand function does not necessarily include a variable that characterizes uncertainty, per se. If that probability density function had included a "conditioning" variable, such as weather, that variable would have been included in the demand function. However, the range of that function in this case, are included.

In the case illustrated, quality uncertainty between L and H affects the shape of the demand function but does not affect the list of "shift variables" included in the implicit function describing Al's demand for good A—although changes in either the utility function (tastes) or the frequency distribution of quality would affect the shape of the demand function and affect the quantity of the good demanded.

Note that the domain of the integral is determined by the probability density function. In the case used above, there is presumed to be a lowest (L) and a highest (H) quality. In other cases, such as the normal distribution, the limits would minus infinity and plus infinity. Some density functions are "full domain." Anything may happen, but some events are extremely unlikely.

One uses the term "probability density function" (pdf) rather than "probability function" here, because probabilities are associated with integrals of (areas under) the density function. Thus, the total area under both a conditional and unconditional probability function is 1 (by definition).

III. Risk Aversion and the Demand for Insurance

DEF: An individual is said to be *risk averse* if the expected utility of some gamble or risk is less than the utility generated at the expected value (mean) of the variable that determines utility (here V).

A **risk averse** person is one for whom the expected utility of a gamble (risky situation) is less than utility of the expected (mean) outcome, if the latter could be obtained with certainty. In

---

[1] If one characterizes a utility function with a concrete function, then these integrals can often be evaluated. This is also possible for single-variable abstract functions—integrals of which simply return the initial integrand, which are then evaluated at the high end of the range of possibilities (here H) and then the low end of the possibilities (here L), which is subtracted from the high end value, which in turn gives one the (net) marginal utility of a change in Q.

mathematical terms, a person is risk averse if and only if $U(x)^e < U(x^e)$ where X is a binary random event, with one possibility, x', occurring with probability P and the other occurring with probability (1-P). $x^e = Px' + (1-P)x''$. This property is true of every possible pair of stochastic outcomes for a risk averse person. **This property implies that any net benefit or utility function that is strictly concave with respect to income, exhibits risk aversion**.

Why? Because expected utilities are convex combinations of utilities. Recall that a function is strictly concave only if af(x') + (1-a)f(x") < f(ax' + (1-a)x") for any x' and x" and any value of a with 0<a<1. If one substitutes a probability for the term a, you can see that the two definitions are essentially identical.

A ***risk neutral*** individual is one for whom the expected utility of a gamble (risky situation) and utility of the expected (mean) outcome are the same. $U(x)^e = U(x^e)$. A ***risk preferring*** individual is one for whom the expected utility of a gamble is greater than the utility of the expected (mean) outcome. $U(x)^e > U(x^e)$.

The degree of risk aversion is often measured using the *Arrow-Pratt* measure of (absolute) risk aversion: $r(Y) = -(d^2U/dY^2)/(\frac{dU}{dY})$ which is a measure of how steeply downward sloping the marginal utility of income is at a particular point. In general, this implies that the more steeply downward sloping the marginal utility of income curve is, the more risk averse an individual is. (In the illustrating example above where $U = V^{.5}$, the marginal utility function is quite steeply downward sloping $\frac{dU}{dY} = .5V^{-0.5}$ and $\frac{d^2U}{dY^2} = -.25V^{-1.5}$, so

$$r(V) = -\frac{[-.25V^{-1.5}]}{.5V^{-0.5}} = (.25V^{-1.5})(.5V^{0.5}) = .125/V$$

The utility Functions that imply risk-averse behavior are all strictly concave, as illustrated below. If the above individual, Al had been risk neutral, the Arrow-Pratt measure would have been zero. If $\mathrm{U} = \mathrm{V}$, then $\frac{dU}{dY} = 1$ and $\frac{d^2U}{dY^2} = 0$, which implies that $r(V) = \frac{0}{1} = 0$.

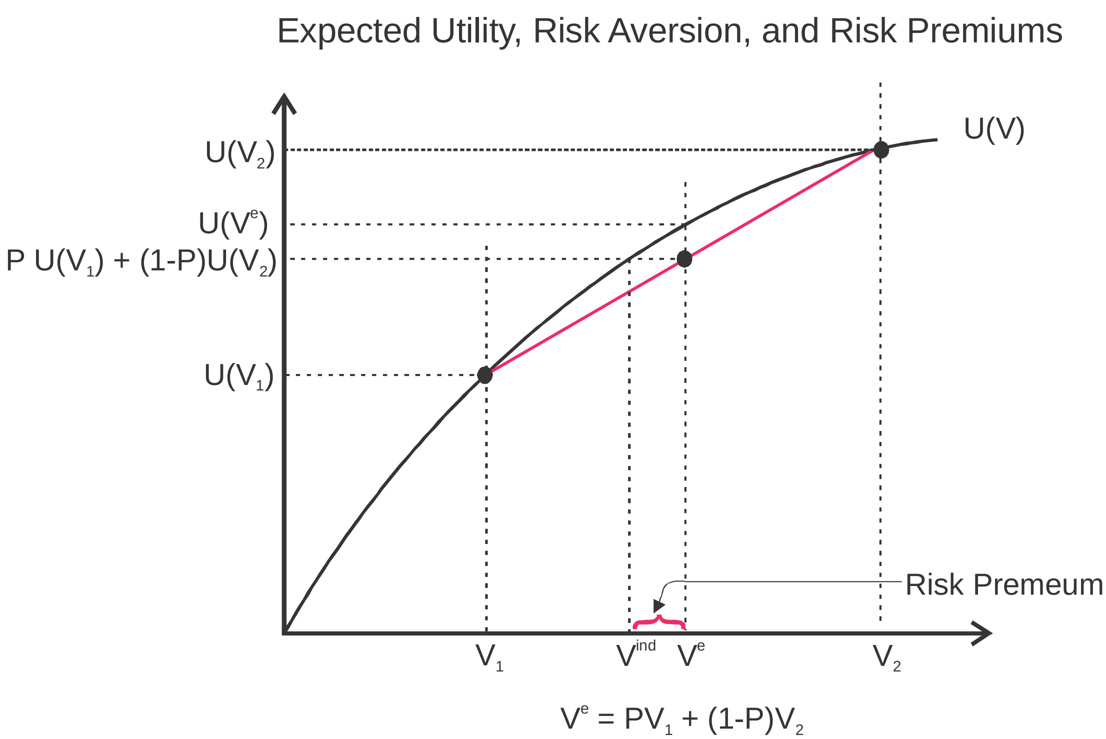
The Geometry of Risk Aversion and Risk Premia

The figure below illustrates a choice setting in which an individual is risk averse and facing a risky environment in which either an outcome with the value V_1 or another outcome with the value V_2 will occur. The individual cannot influence which outcome it will be, but knows that the probability of V_1 is P, which implies that the probability that V_2 is (1-P) (Recall that that the probabilities for the only two possible events have to add up to one.)

Let's refer to the individual as Al. Al's utility function is strictly concave, which means that a cord connecting any two points on it lies below the utility function (except for the two points used as end points—which, by definition, are not part of the cord). Assume that Al confronts an uncertain environment in which V_1 occurs with probability P and V_2 occurs with probability (1-P) Al's expected utility in that case is:

$$U^e = PU(V_1) + (1-P)U(V_2)$$

As P increases from 0 to 1, the expected utilities trace out the cord between $U(V_1)$ and $U(V_2)$ and so will be below the utility function if it is strictly concave.[2]



[2] We have until this point used a "sufficient condition" for strict concavity, namely that a utility function is strictly concave if it has a positive first derivative for V and a negative second derivative for V. In other words, Al's utility function is concave if it exhibits diminishing marginal returns from V. However, however at this point the formal definition becomes a nice bridge between risk aversion and strict concavity. The expected value of V is $V^e = PV_1 + (1-P)V_2$. Note that if U is strictly concave then $U(V^e) > PU(V_1) + (1-P)U(V_2)$. As mentioned above, this looks exactly like the definition for concavity except that we've substitute "P" for "α".

This geometry is illustrated in the diagram below for a probability, P, that is approximately equal to 0.5, but it would be true for all probabilities $0< P < 1$ and all strictly concave utility functions.s

This diagram can also be used to determine how much an individual would be willing to pay to have a certain payoff rather than face a risky or uncertain future. This is done by looking at the certain outcome that a person would be equivalent in their mind to the risky event.

If we go to the left from the expected utility associated the two probabilistic outcomes over to the utility function and then down to the horizontal axis, we find the value (labeled V^{ind}) that Al would find equivalent to the risky one faced. (V^{ind} is the certain outcome that generates the same expected utility as the risky one faced.) The difference in values, $V^e – V^{ind}$, (assuming that the values along the horizontal axis are in money terms) is the highest price that Al would pay to avoid the risk.

It is also the lowest value that Al would accept to bear the risky environment shown rather than have outcome V^{ind} with certainty. That difference is called Al's **risk premium** for this choice setting or "gamble." Al would accept the gamble (risky environment) rather than V^{ind} only if the expected value of the payoff is at least that much greater than V^{ind}.

Very similar logic can be used to determine the highest amount that Al is willing to pay for insurance. Note that the expected loss can be represented as $P(V_2\text{-}V_1)$ which is the distance from V_2 to $V^e=PX_1+(1\text{-}P)V_2 = P(X_1 – V_2) + V_2 = V_2 – P(V_2\text{-}V_1)$. This last expression characterizes the expected value of the risky setting in terms of the loss that occurs when the unfortunate event occurs—possibly a fire, accident, or a disease. The risk premium is the amount above the objective risk that Al is willing to pay to avoid the risk.

IV. Applications

1) Selling Fire Insurance

The existence of risk premiums plus the effects of sample size on sample means, thus, implies that selling insurance can be profitable. In our example, fire insurance transfers risk from homeowners to insurance companies. However, if the probability function is well known and the insurance company has many customers, the insurance company has only a very small risk. The mean payout from selling insurance would be approximately $P(V_2\text{-}V_1)$ per customer, per year, and the price for the insurance can be **up to $P(V_2\text{-}V_1)$ + the risk premium** from the above figure. This implies that selling insurance can be profitable—although it does not guarantee it.

If a large number of purchasers for an insurance product exists, firms will have a quite predictable flow of expenses that are approximately equal to the expected value of the average loss, while customers are willing to pay more than that to avoid the risk of such losses. If the risk premium customers are willing to pay is more than enough to cover the cost of sales and administration of the insurance products sold, and insurance companies may be profitable investments. However, competition among insurance providers, in turn, tends to bring profit rate down to the "ordinary" rate of return that firm owners make from their other investments. But the prices for insurance in equilibrium will be sufficient to keep them in business.

Note that both the buyer choices and the insurance provider choices involve the mathematics of expected values rather than utility or profit maximization under certainty. Insurance is a market that would not exist without measurable risks and the ability to shift risk from one group to another. The ability to moderate risks through risk pooling—which is a property of the variance of sample averages, which tend to fall as sample size increases—is another key feature of the risks for which insurance is possible. (Recall that the variance of a sample mean is var(sample mean) = σ^2/n, where σ^2 is the variance of the variable being sampled and n is the sample size.)

2) Application: Expected Benefits Maximization and Uncertain Product Quality

Another case in which probabilistic thinking is likely to take place is regarding products of uncertain quality at the point of sale. For example, suppose that Al is considering purchasing some produce from a farm and knows that some of the produce will be of high quality (H) and some will be of low quality (L) but simply can't tell the difference between the two types or produce at the time of purchase, as is true of many types of produce (corn, potatoes, tomatoes, squash, etc). However, suppose that Al has sufficient experience with the farm or produce shop to know what the probability of a defective product is.

Suppose that there are just two levels of quality that tend to turn up, high quality and low quality. Suppose also that the probability of high quality is F and that price per unit is simply P. Suppose that the benefits of high-quality units is B(Q, H) and the benefits from quality units is B(Q, L) where B(Q,H) > B(Q,L) for every Q.

How many units will Al purchase? Al's expected net benefit from purchasing produce is expected benefits less expected costs:

$N^e = FB(Q,H) + (1-F)B(Q,L) - PQ$

To find Q*, differentiate Ne with respect to Q and set the result equal to zero.

$F (dB^H/dQ + (1-F)(dB^L/dQ) – P = 0$

The first two terms of the expected marginal benefit of the produce and the last is its marginal cost. To find a specific value we would need to use concrete functional forms for the two benefit functions, as with $B^H = HQ^{.5}$. and $B^L = LQ^{.5}$, with H>L, in which case our first order condition would be:

$.5FH/Q^{.5} + .5(1-F)L/Q^{.5} = P$

Multiplying both sides by $2Q^{.5}$ yields $HF + L(1-F) = 2PQ^{.5}$, which implies that $Q^* = [HF + L(1-F)]^2/4P^2$

In either case, the quantity Al purchases rises with F (the probability of the high-quality type) and with the benefit of the high-quality product, H (an indication of the quality of the high-quality type) and falls as low quality version of the produce decreases or price increases.

This risk is potentially insurable, but it may be too difficult to organize transaction by transaction and so remains an unprofitable type of insurance. (Notice the money-back-guarantees may reduce the buyer's risk, shifting it to the seller—who may then increase his or her price by more than the expected loss per customer. In effect, such guarantees make the seller an insurance company, and they can charge a premium for that service that their risk-averse customers are very willing to pay (at least up to their risk premium).

3) Applications: Quality Control—A Role for Management and Monitoring

Of course, quality variation is not only associated with agricultural products. All good and services have some variation in quality. Within mechanized construction processes, wear and tear, and product failures generate variation in quality affecting the usability of the products produce for buyers of the product. Such variation may affect the durability of the product sold as well as the benefits that it provides to buyers.

Insofar as quality variability can be estimated by consumers, and purchase decision are essentially independent of one another, we can use the net benefit maximizing model to characterize the demand for such products. $N^e = B^e(Q) - C(Q)$ which can be represented as $N^e = Fb^L(Q) + (1-F)b^H(Q) - PQ$ for the two-quality case, where F is the relative frequency of low-quality units and (-F) is the relative frequency of high value units and P is the price of the units purchased. The benefit functions are both assumed to be strictly concave.

The quantity that a purchaser would acquire would satisfy the first order condition: $dN^e/dQ = Fdb^L/dQ + (1\text{-}F)db^H/dQ - P = 0 \equiv H$. The implicit function theorem implies that $Q^* = f(P, F)$ with $dQ^*/dP = dH/dP/\text{-}dH/dQ = (-1) / -[Fd^2b^L/dQ^2 + (1\text{-}F)d^2b^H/dQ^2] < 0$ given the strict concavity of the benefit functions. The demand function is downward sloping in price. Similarly, $dQ^*/dF = dH/dF/\text{-}dH/dQ = [db^L/dQ\text{-}db^H/dQ] / -[Fd^2b^L/dQ^2 + (1\text{-}F)d^2b^H/dQ^2] < 0$. As quality diminishes, demand falls as well.

Notice that a firm selling this product and facing a downward sloping demand curve can influence the extent of demand though decision that affect the frequency of low-quality units. For example, a firm's monitoring expenditures, M, may reduce F, with $F = h(M)$. In effect firms have two controls in their efforts to maximize profits—monitoring (M) and output levels (Q), rather than simply one as usually assumed. Moreover, in this case, the demand curve faced is partly determined by the firm's decision about quality.

Let $P = g(Q, h(M))$ characterize its inverse demand function. The firm's profits in this case can be characterized as $\Pi = g(Q,h(M))Q - c(Q,M)$ (assuming that input prices are constant in the period of interest). There will be two first order conditions for its profit maximizing efforts:

$$(dg/dQ)(Q+1) - dc/dQ = 0$$

$$(dg/dF)(dF/dM)Q - dc/dM = 0$$

Both first order conditions are simultaneously satisfied at the firms profit-maximizing output. This makes them a bit difficult to characterize in words. The monitoring decision will affect output decision by shifting the demand curve. And the output decision will affect monitoring by affecting the extent to which prices are affected by monitoring (dg/dF). But in each case, the ideal levels occur where marginal revenue generated by changing Q or M equals the marginal cost of Q or M.

Notice that this relationship makes perfect sense in any market where firms face downward sloping demand curves, but is makes less sense in settings where the firms face are price takers and face horizontal demand curves. An increase in quality in that case, might cause all consumers to want to purchase the product from the sole firm that has successfully improved its quality, which it cannot do. Or, it may cause the market to at least temporarily split in half, with one group of firms proving lower quality outputs than the other.

V.Some Additional Applications of expected utility and expected profit maximizing choices

The 1960s was a period in which rational choice models were applied to fields generally regarded by most economists to lie outside of economics. These new areas of research gradually gained sway inside economics and expended the field to areas of the economics of regulation, socio-economics, law and economics, and political economy. Several of these areas made use of the expected utility maximizing model and expected net benefit maximizing model. Several Nobel prizes were awarded to the pioneers in these new areas of research, several of which are taken up in part III of this book.

4) Applications: Expected Values and the Effects of Regulation

i. One can also use this type of model to model the effects of economic regulation.
   - For example, in the area of environmental regulations, firms will take account of their overall net benefits from pollution including both cost savings and anticipated regulatory fines when choosing their production methods.
   - In the absence of fines or fees for pollution and in the absence of enforcement of fines greater than 0, firms will choose their production methods to minimize their production costs—as in the models developed in the first part of the course (prior to the midterm).
   - (This does not necessarily mean that firms will pay no attention to air or water pollution, but they will do so only insofar as it affects the firm's expected profit through productivity and cost effects. Air or water quality that *affects the productivity of the firm's workforce* will be taken account of, but not spillovers on others outside the firm.)

ii. In the real world, regulations are only imperfectly enforced, and firms know this.
   - Consequently, it is not simply the magnitude of the fine or penalty schedule that affects a firm's decision to "pollute illegally or not," but also the probability that a person that violates the law will be caught, convicted and punished.
   - Analyzing regulatory law and its enforcement on a firm's choice of production method and output level requires taking account of both the "expected cost" and "expected marginal cost" of any fines or penalties that might be associated with its production and output decisions.
   - (In addition, firms might face a loss of reputation and therefore reduced demand for their products if they are found guilty of violating regulatory law, but that effect will be ignored or assumed to be part of the fine.)

iii. Consider a case in which production methods are fixed and output is regulated—which is the easiest case to model.

- **In a regulatory environment with fines, a pragmatic firm's expected profits equal its total revenues less its production costs less its expected fines**: $\Pi = R - C - F^e$ where $F^e = PF$
- Suppose that Acme's output is sold in a competitive market, its cost function is $C=cQ^2wr$ and that its expected fine is the probability of being caught and convicted, which increases with output in excess of the regulatory limit, $p(Q-Q^R)$ and a fine schedule that increases with the extent of the violation $f(Q-Q^R)$ for $Q> Q^R$.
- $\Pi^e = PQ - cQ^2wr - p(Q-Q^R)f(Q-Q^R)$
- To make the functional form a bit more concrete, let us assume that $P(Q-Q^R) = a(Q-Q^R)$ and $f(Q-Q^R) = b(Q-Q^R)$. In this case, Acme's expected profits are:
- $\Pi^e = PQ - cQ^2wr - a(Q^R-Q)\, b(Q^R-Q) = PQ - cQ^2wr - ab(Q-Q^R)^2$
- Assume that the regulatory constraint is binding on Acme, and so it will take the expected fine schedule into account when making its output decision. Its expected profit maximizing output can be characterized by differentiating the above function with respect to Q, which is a bit more complex than usual because of the "$Q-Q^R$" terms.
- $\Pi^e_Q = P - 2cQwr - 2ab(Q-Q^R) = 0$

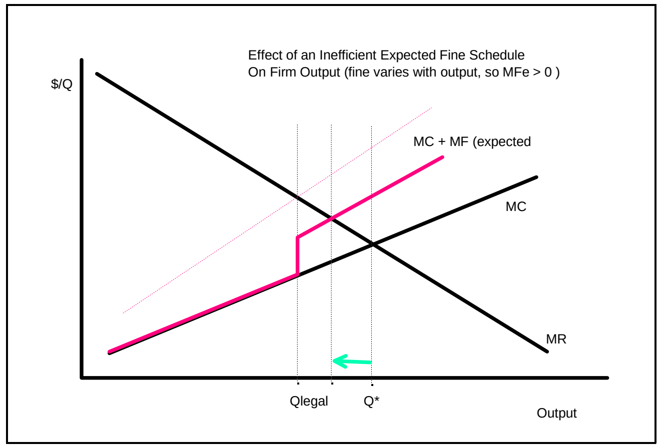
iv. This can be solved for Q^* . First, shift the Q terms to the left side of the equal sign:

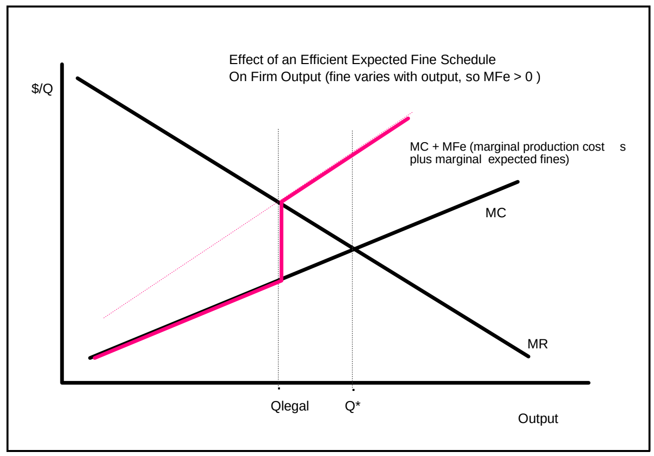
- $P = 2cQwr + 2ab(Q-Q^R) = Q(2cwr + 2ab) - 2abQ^R$
- Adding $2abQ^R$ and dividing yields:
- **$Q^* = (P+2abQ^R)/(2cwr + 2ab)$**
- **This is Acme's supply function in the regulatory environment modeled.**

v. Note that its output now varies with the regulatory standard (Q^R) its input costs (w and r) and parameters of the probability of being fined and fine schedules (a and b).

- Acme's output declines as input prices and the expected fines increase (w, r, a, or b increase) and increases as the regulatory threshold (Q^R) increases.
-
- (Another possible output is simply Q^R, but this cannot be modeled with calculus because of a discontinuity in the expected cost function at that quantity. See below.)

A. The diagram to the left illustrates Acme's decision in this type of setting (with somewhat simpler probability and fine schedules).





B. For students that have had public economics, note the similarities between Pigovian taxes and optimal enforcement with fines.

   If the regulation attempts solve an externality problem and achieve Pareto efficiency, Q**, then the smallest fine sufficient to induce the target Q** has the **same expected value** as a Pigovian tax at Q** (with $Q^R \leq Q^{**}$). The expected fine should equal the expected marginal damages done by the Q**th unit of output.

C. Note that there is always a policy-tradeoff between the probability of conviction and the optimal level of punishment. [Recall that the expected fine is $F^e = PF$]

D. Some General Implications

   i. The larger the fine, the smaller the probability of capture can be to generate the same effect on individuals.

   ii. The larger is the probability the smaller the fine can be and still have the same effect.

      a. The effect is determined by the expected fine, PF, in this case.

      b. The probability that an illegal activity is detected and punished varies with the resources used to enforce the law and the flagrancy of the violation, so the probability of being caught and punished tends to vary with law enforcement budgets and the size of the violation.

      c. The politics of enforcement and penalties are partly determined by error rates in detecting criminal activities--sometimes the wrong person is singled out for punishment.

- Puzzle. Given this, how would you pick the appropriate punishment for speeding? for theft? For murder? etc.
- Puzzle. How would the relative importance of the probability of detection and the expected fine be affected by the process of a jury trial and a long delay between being detected and being fined? (Some ideas for doing so are provided in the next chapter.)

- Puzzle: Write down an expected profit function for a firm facing a fine schedule that is imperfectly enforced, but where the fine increases as Q exceeds Q legal. Find the first order conditions and compare them to the above diagrams.
- Puzzle: draw examples of a perfectly enforced and imperfectly enforced "fixed fine schedule." (Such fines do **not** affect expected marginal costs.) Compare your graph with the mathematics of expected profit maximization in this case. Are such fines always irrelevant?

5) Applications: Expected Values and the Logic of Crime and Punishment

i. The economic analysis of crime derives from a classic paper written by Gary Becker (1968), who subsequently won a Nobel prize in economics, only partly for that contribution. In that paper, and in many others published since then, a criminal is modeled as a rational agent interested in maximizing his EXPECTED income or utility, given some probability of punishment.

ii. This type of model can be used to model theft and violations of other laws.

- In the real world, criminal laws are only imperfectly enforced, and both criminals and ordinary persons who occasionally think about violating a law or two know this.
- For example, a net income maximizing criminal would maximize an expected function like $\Pi^e = PQ - cQ^2 - p(Q)F$ where Q is the number of crimes (thefts), price is the average price received by "fencing" the stolen goods, p(Q) is a probability function describing the way that that the probability of being caught and convicted varies with the number of crimes and F is the financial penalty assessed (or if jail time is spent, the opportunity cost of the time spent in jail and any subsequent losses in earnings).
- The rational theft chooses Q* such that $\Pi^e{}_Q = 0$, which in this case requires Q* to satisfy $P-2cQ - p_QF = 0$ or $P = 2cQ + p_QF$ (set the marginal revenue from theft equal to its **expected marginal cost**, which is not known with certainty).
- Let's give the probability function a concrete form as with: $p = aQ_2$ then $p_Q = 2aQ$ and the above first order condition becomes $P-2cQ - 2aQF = 0$ or $P= 2cQ + 2aQF$, which can be solved for Q.
- $P = Q(2c + 2aF) \rightarrow$ **$Q^* = P/(2c + 2aF)$**
- Note that this implies that the rational criminal responds to incentives, his or her crime rate falls as the probability of being caught and convicted rises (e.g., with 2a), as the fine increases, and as the marginal cost of theft increases.
- Note also that there are tradeoffs between the size of the fine and the probability that a criminal is caught in terms of their overall effect on the criminal.
- (This model provides a short form of Gary Becker's classic 1968 paper on crime and punishment.)

iii. Many other examples from law and economics can also similarly modelled. One does not have to be a more or less professional criminal for this logic to apply.

- One can think of choices to drive faster than the speed limit on a highway or to park without putting money in a parking meter, or to trespass on a neighbor's property, fail to report some income on one's taxes, and so on in much the same manner.

VI. Some Conclusions

The main implication of this chapter is that neoclassical economics and its associated models can be easily extended to take account of risk—that is to say, choice settings in which outcomes are uncertain, but the probability of various outcomes can be accurately (or exactly) estimated. In such case, the logic of optimizing choice applies and the choices can be modelled in the usual way using diagrams in calculus that take account of risk.

There are some new implications.

First, that individuals with similar tastes may nonetheless differ in their degree of risk aversion. Individuals may all prefer more income or wealth to less income or wealth, but their utility functions (to the extent that these can be worked out or estimated) may differ in their curvature, in their degree of strict concavity. The Arrow-Pratt measure of risk aversion is one way to measure such differences. Some people are more risk averse than others, and so willing to pay more for insurance than others.

Second, risk aversion creates markets for insurance. If everyone were risk neutral, insurance like products might still exist, but there would be no sales of such products, because consumers would not be willing to pay a premium for those products—and that premium is necessary to cover the cost of administering insurance products.

Third, insurance products and similar products not always considered to be insurance imply that risks can be shifted from one individual to another and from one group to another. Moreover, they also demonstrate how risks can be pooled by insurance companies in a way that actually reduces overall risks, because of the statistics of sample means. As the sample size increases, the variance of the sample mean falls, and thus risk also falls. Insurance companies with accurate estimates of the probability of an insurable event, know almost precisely the amount that they will pay out from year to year.

Moreover, competition among insurers—as noted by Frank Knight—tends to reduce insurance company profits (for honest companies) to "ordinary" rates of return—the so-called zero-profit equilibrium of Marshallian perfect competition. It may not do so entirely, because of differences in size and organization, but the Marshallian implications can serve as a first approximation when there are large numbers of firms that provide insurance.

Fourth, there are many situations in which risks are not pooled and in which consumers, firms, or entrepreneurs have to make decisions in risky choice settings—shall I bring an umbrella or not? In these cases, one can use expected utility or expected profit maximization to model the decisions reached and their comparative statics. Although experimental evidence suggests that individuals do not perfectly behave as predicted, in most cases, the models provide a good first approximation for the average behavior and for how changes in circumstances affect that behavior.[3]

[3] See, for example, for an overview of prospect theory see KD Edwards (1996) or Tversky and Kahneman (1992, 2013)—who won a Nobel prize for their work in this area and others, that show some limits of the rational choice model.