Chapter 9: The Economics of Information

I. Uncertainty and Information

The last two chapters have explored implications of choice settings in which individuals and organizations are uncertain about the consequences associated with their actions. The consequences of an action or purchase may be probabilistic—as when one purchases a lottery ticket or a product whose quality is characterized by a probability function such as the uniform or normal distribution. Or the choice setting may be one in which the consequences of an action or purchase are (or may be) results of such complex or chaotic processes that no probabilistic representation is possible. In the latter case, individuals may nonetheless use probabilistic ideas to try to assess the relative merits of alternatives using subjective (Bayesian) rather than objective frequentist characterizations of probability functions. Such subjective assessments, for example, allowed us to use expected utility and expected profit models to characterize the research and development activities of firms. Probabilistic thinking—even when no specific probability function is known beforehand—allows individuals to be systematic in making choices in settings of uncertainty given what is known beforehand, although it does not eliminate mistakes. Individuals and organizations may later regret such decisions, because the probabilities used are just guesses, rather than rooted in observation.

It is important to realize that both choice settings, individuals have been assumed to know as much as it is possible to know. If a process is truly probabilistic, one can only know the probability function, not the outcome of such processes. Similarly, if a process is truly chaotic, little systematic can be known about it beyond the fact that it falls into that category of phenomena.

Essentially all of the models developed to this point in the text have assumed that individuals know as much as can be known about their choice settings. As a consequence, the analysis has not considered the possibility that additional information may change the nature of the choice problem setting confronted. That is to say, useful learning has been ruled out—or rather, set aside—to facilitate the analyses undertaken in previous chapters. This chapter explores how additional information (learning) can affect an individual's or firm's assessments of the possibilities confronted and their consequences.

In some cases, additional information "simply" reduces the variance of the probability functions used by decision makers. Information of this type can improve one's decisions by improving one's estimates of parameters of relevant functions such as a probability function's conditioning variables.¹ In other cases, new information may reveal possibilities or causal chains that had not been previously imagined or taken into account. For example, the new information provided by satellites have made short-term weather forecasts much more reliable than they were a century ago—not simply be improving our estimates of well-understood probability functions but by adding previously unknown causal factors that tend to produce one pattern of weather rather than another.²

There are areas of knowledge where new insights are unlikely, such as lens optics, because received wisdom about particular processes through which lenses operate have proven to be accurate in many circumstances. However, new insights are commonplace in many areas of knowledge. Economics, for example, remains an active field or research partly for that reason.

Moreover, new insights are far more common at the individual level of analysis than at the level of humanity, because none of us has time to master all that is known by humanity at any given moment. Thus, there are many choice settings in which additional information can improve an individual's choices by improving ones understanding of familiar phenomena or by unveiling entirely new phenomena or relationships that create new opportunities that can be taken advantage of or reveal new risks to be mitigated. In both types of choice settings, additional information is useful, which creates a demand for information. As a consequence, individuals and organizations are constantly gathering and interpreting information. Living is substantially about learning—which is true of economic and other areas of life..

Even when we walk across a well-marked pedestrian crossing of a paved road, rather than decide where each foot should be placed before crossing the street and then closing our eyes and

¹ The idea of a conditioning variable for a probability function is not obscure—although we have not used that idea before. An example of a conditioning variable is cloud cover. When the sun is shining it is less likely to rain than when dense (dark) clouds are overhead. Conditioning variables are often written as p(x) = f(x | z) where p(x) is the conditional probability function, f is the conditional probability, and z is the conditioning variable. The conditional probability of rain p(x) varies with the degree and darkness of the clouds, f(x, z) in the example just used, where z is he cloud density.

² An example of this is the effect of El Niño conditions in the Pacific Ocean west of South America on weather patterns in North America.

walking as planned, we take a step or two and then consider where to place the next few steps based on fresh information about the road surface (potholes or trash), the apparent path of other pedestrians, and our balance after we have taken the previous step. As a consequence, most of us continually adjust our planned path as we walk across a street in light of new information. The same is true of many aspects of one's personal life—including many that are relevant for market prices, outputs, and patterns of exchange.

The Demand and Supply of Information

Some general aspects of the demand and supply of information are relatively easy to analyze conceptually—although particular instances of learning are not easy to carefully model. Acquiring and processing information (data) takes time and attention—both of which are scarce resources. As a consequence, we all make decisions about the types of information that we collect and analyze that resemble choices among products. Our analysis produces new information for ourselves and often information that would be valued by others. Which is to say, our "processing" of new information in light of other information previously possessed always increase's humanity's supply of information, at least a bit. Thus, the demand for and supply of information are closely connected—more closely than are the supply and demand for other goods and services.

That we all must allocate our scares resources (time and attention) among information sources allows us, at least in principle, to use models similar to those used to model our choices of expenditures on products sold in markets.

Producing information also requires time and attention (and in some cases special tools for observing and measuring the phenomena of interest). Thus, it looks quite a bit like a firm's normal profit maximizing choice. An information producing organization allocates time, attention, and other resources to produce the most information possible from its teams and laboratories. This process investigates phenomena, develop and test possible explanations, faithfully records the results, and disseminate the resulting information through conversations, media of various kinds, some of which is sold as "expert opinion" to individuals and firms with demands for the kinds of information produced.

In cases in which information is produced with profits in mind, and thus the media in which the information produced is disseminated are priced and sold in markets—as with newspapers, some websites, and courses in most colleges—prices adjust to set the quantity demanded equal to the quantity supplied. In such cases, the models of Part I of this textbook shed useful light on the types of information produced and sold in markets. Other information is given away freely, which requires other explanations.

However, there are relevant differences between markets for various kinds of information and markets for other goods and services. For example, neither consumers nor producers of information precisely know the value of the information collected and analyzed until after it is collected and analyzed, because new information—more or less by definition—cannot be fully known until it is produced.

There are many books in university libraries (and now on the web) that an individual can read and many skills that we might master. We do not have the time or attention to read or master them all, so we make choices about which books to read and which skills to attempt to master. However, exactly what one will learn by reading a book or taking a course remains unknown until the book is read or course is taken. And, similarly, the authors of those books rarely knew exactly what would be in them until the final drafts of their manuscripts were written.

Relevance of Learning for Markets

A large subset of these informational choices and resultant learning are relevant for understanding how markets operate. The most commonplace of these effects are the main focus of this chapter.

For example, consumer knowledge affects the prices paid for goods and services. No consumer knows all the prices of all the good sold in markets, as assumed in previous chapters. Consumer knowledge and ignorance thus affects the kinds of goods and services demanded, and, indirectly, the kinds of goods and services produced and brought to market for sale. It does so because information about consumer wants obviously inform the output decisions of firms (a subject taken up in more detail in Chapter 10).

Information also affects the degree to which individuals and organizations specialize, and thereby the manner in which the things brough to market are produced. Specialization among humans is mostly an informational phenomena. We know some things very well (our specialities) but are largely ignorant of many others. Furthermore, product innovation is possible only because all possibilities are not fully understood or taken advantage of. Hayek (1945) once argued that it is differences in the knowledge possessed by individuals that makes market prices and specialization critical determinants of prosperity.

Thus, understanding how information affects market-relevant decisions is central to understanding how markets operate.³

II. Modeling the Demand for Information

The demand for information differs from the demand for most goods and services because it is a demand for something that cannot be known beforehand. If one already knows a price, probability, or other fact, then it is unnecessary to spend one's scarce time and attention acquiring such information. It is only in areas in which one is less informed than one believes would be useful that it can be worthwhile to devote part of one's scarce time and attention to acquiring additional information or to undertaking additional processing of information that one already has.

Many choices about information are instances of normal risk taking in that one has reasonable expectations about distribution of possible results associated with gathering more information. For example, the current price of a good or investment may be relevant for some other choice, and being familiar with the past values of the price of interest, one knows quite a bit about the probability that the price will take various values through time. Knowing the current price, however, may be particularly relevant for a buy or sell decision of a stock or for the purchase of a particular good—is it "on sale" or not? In other cases, the expectations may be less concrete, but still reasonable. Students that have taken several economics courses may confidently predict that "If I take more economics courses, I'll know more about economics afterwards than when I started" without knowing exactly what will be learned.

³ Some individuals may not be aware of the scope of their ignorance—in effect, they are ignorant of their own ignorance—and so act with what Adam Smith once termed "over-weening conceit." However, the attitudes of "know it all's," while comforting, tends to make such decision makers more error prone than they could be. It tends to make them more confident than they should be, and thus more likely to be risk takers than they should be. The latter leads to many bankruptcies and failures of new firms. However, it should be noted that their occasional surprising successes often benefit the more sensible prudent men and woman who more accurately assess the riskiness of such ventures, and so would never have attempted to develop the unlikely—but sometimes very successful—products, production methods, and economic organizations created by such overconfident entrepreneurs.

Others informational choices resemble Knightian uncertainty. In cases of complete ignorance or in which one has only a bit of information about the phenomena of interest, one lacks sufficient knowledge to construct an observation-based probability function. One may make guesses about what one will know after such an investment in information is made, but prior to the investment, one does not really know what will be learned. A student that takes his or her very first economics courses may have no idea what will actually be learned—and only realize well after the course was taken whether it was worthwhile or not. The information acquired in such cases may be a complete surprise!

In both settings, choices are made to acquire additional information and in both cases, the information obtained may improve subsequent decisions—but the two types of decisions are quite different from one another. In the first case, more or less rational decisions can be made that can be characterized with the rational choice models developed in this textbook (and in economics and game theory more generally), in the second, such calculations cannot really be undertaken. Although as mentioned before, one can use the expected utility and expected profit approach to gain some insight into the factors that affect such decisions, those models cannot fully capture Knightian uncertainty or much about the individual calculations that induce them to seek such information.⁴

In many cases, the collection and processing of information reduces uncertainty and thereby makes decisionmaking less error prone. However, there are also cases in which more information reveals uncertainties than one did not knew about beforehand, which increases uncertainty. Such learning can induce better planning not by reducing uncertainty, but by better accounting for risks that previously were unknown.⁵

⁴ Again, this chapter is not attempting to characterize the demand for information as entertainment—which is an important rationale for collecting information characterized by Knightian uncertainty. Tourists and scholars, for example, often put themselves into unfamiliar contexts so that they will be surprised (hopefully more often pleasantly than unpleasantly) by the information (experiences) acquired. It is the discovery of the new that motivates the collection of such information, rather than any clear idea of improving one's future decision or earning profits by selling it.

⁵ Accumulating more information does not always reduce uncertainties, as some statistical measures of information imply, but nonetheless may improve decisions by providing a more realistic understanding of a choice setting. This is a flaw in many statistical measures of information, which were initially developed as models of communication. Improved communication occurs when errors in interpretating the messages sent diminish. Such communication is said to be more informative. See, for example, Shannon's classic book (1948) or Shannon and Weaver (1998). This usage of the

A. The "Search" Model of the Demand for Information

We begin with one of the simplest models of the demand for information and its associated learning. It is a model that is based on the "search" model introduced by George Stigler (1966) which attempted to analyze the extent to which informational costs could account for the dispersion of prices in real markets.⁶ Outside of organized stock and futures markets, there is usually more than one price for goods sold in markets. Prices do not converge to a single price but to a cluster of prices around some average price—an average which may be regarded as the market clearing price of the models developed in the first part of this book. There are other explanations for price variation—for example differences in the "free" services or waiting times offered by different merchants—but the informational explanation is important and emerges from a straightforward extension of the rational choice model. It turns out that the prices that one pays for goods are partly a matter of choice rather than entirely beyond by market forces beyond one's control.

In the leanest neoclassical models, all knowledge problems and transactions costs are assumed away, in which case the 'law of one price' holds. If consumers know the prices of every good on offer, they will always purchase goods at the lowest available price. Thus, only one price can exist, namely the lowest one sufficient to keep suppliers in business. The lowest one that can be sustained is the one that equals the average cost of production (for the marginal firm) as pointed out by Marshall in the late nineteenth century among many others. Once the lowest sustainable selling price emerges, a price lower than that price would attract all the buyers in the market, but firms charging that lower price will lose money and go out of business. Any price higher than the minimum on offer will attract no customers—since there is no reason to spend more on an identical product than the lowest price available. (Differences in services and/or locations in neoclassical models imply that products are different, rather than being identical products. It is the full package of goods, services, and location that identify a "product.") It is this type of logic that was implicitly relied on when the neoclassical theory of equilibrium prices was reviewed in chapter 5. Unique equilibrium prices emerge in such models.

term information has by now produced a very large and technical field of research. Most of that field, however, is beyond the scope of this chapter, which attempts to model some of the essential features of more or less informed decisions, rather than communication, signaling, or artificial intelligence.

⁶ This paper was mentioned in the mentioned in Stigler's 1982 Nobel Prize. https://www.nobelprize.org/prizes/economic-sciences/1982/press-release/

The logic of "perfect competition" is clear---but what if information about prices is not freely available to all consumers? Stigler (1966) argues that in such cases consumers would sample prices and purchase goods at the lowest price in their sample. That is to say, they would search for the lowest price. However, different consumers would choose different samples, and so would purchase at different prices because the lowest price in their samples would differ. The cost of searching for the lowest price and the expected savings on purchases would determine how much searching an individual undertakes.

Stigler's theory has often referred to as the search theory of information, but it is better thought of as the sampling theory of prices. He argues that the easier it is to search for prices (e.g. the less costly sampling is) the narrower is the range of prices that tends to emerge in markets for a homogeneous product. Searching (sampling) limits the range of prices paid for goods and services because consumers purchase the good at the lowest price in their sample. Thus, firms with well above average prices sell relatively little of the good or service of interest—e.g. only to those who do very little or no searching. If most people do a bit of shopping around, the result is reduced sales and profits for the higher priced merchants, and greater sales and profits for lower priced firms. Thus, Stigler's theory can account for both the dispersion of prices observed in the real world and for the extent of that dispersion—something the law of one price cannot.

Modelling an Individual's Price Sampling Decision

Suppose that Al knows her own demand function $Q^{D} = q(P, Y)$ but doesn't know the price at which the good of interest will be purchased. However, he or she knows from statistics that as more sampling is undertaken, the lowest price in one's sample tends to fall. In this case, sampling might involve driving to numerous shops and inquiring about the price of the good he or she is thinking about purchasing or simply spending time on the computer looking at alternative websellers for the good of interest. In each case, sampling has an opportunity cost—namely the value of the time spent searching plus any other expenses such as fuel for one's vehicle or electricity and internet service for one's computer or cell phone.

The simplest way to model this choice is as an effort to minimize Al's overall expenditure on the good of interest. Al's expected overall expenditure (E^e) on the good of interest is the amount spent to purchase Q units of the good (PQ) plus the opportunity cost of searching for the best price (C). This can be written as:

$$E^{e} = P(N)q(P(N),Y) + c(N)$$
(9.1)

where N is the number of shops or web-sources sampled, q is Al's demand function and c(N) is the opportunity cost of search—other things being equal. The expected price, P(N), is a function that declines with sample size at a decreasing rate ($P_N < 0$ and $P_{NN} > 0$). The sampling cost function, c(N), increases with the sample size at an increasing rate ($C_N > 0$ and $C_{NN} > 0$). Differentiating equation 9.1 with respect to N characterizes Al's optimal sample size for his or her searching efforts—e.g. efforts to learn what the lowest available price is for the good of interest.

The first order condition looks similar to the usual ones, but in this case, Al is interested in minimizing her total cost for purchasing the good of interest and undertaking search.

$$P_N[Q + PQ_P] + C_N = 0 \equiv H \text{ at } N^*$$

$$\tag{9.2}$$

Or alternatively as:

$$-P_N[Q + PQ_P] = C_N at N^*$$
(9.3)

The first term in equation 9.2 is Al's expected marginal benefit from sampling, which is the marginal reduction in the expenditures on the good associated with lower prices generated by sampling ($P_N < 0$). The second is the marginal cost of sampling. Al's expected marginal benefit from searching varies with the extent to which prices tend to decline with search (P_N) with the quantity likely to be purchased (Q), the price ultimately expected to be paid (P), and the price sensitivity of his or her demand function (Q_P).

Statistical theory implies that as the variance of the price distribution increases, P_N tends to increase, which tends to induce greater search because the marginal benefits of searching increases. As searching for the lowest price increases, this tends to put downward pressure on the firms selling the good of interest at above average prices.

Equation 9.3 expresses the first order condition in somewhat more conventional MB=MC terms. Al's overall expected cost for the good is minimized when the marginal benefits of search (the negative of expected marginal expenditure reduction) equals the marginal cost of searching for a lower price (e.g. the marginal cost of an increase in sampling). Figure 9.1 illustrates equation 9.3.



Comparative statics of this relationship can be characterized using the implicit function differentiation rule. As modeled, Al's income, Y, is the only exogenous variable and it affects demand rather than costs. Thus,

$$N_Y^* = \frac{H_Y}{-H_N} = \frac{P_N[Q_Y + PQ_{PY}]}{-(E^e_{NN})} > 0$$
(9.4)

If an expenditure minimum exists for N>0, then the expenditure function is U-shaped, which implies that an increase in sampling at N* increases marginal expenditures, so $E^{e}_{NN} > 0$ at N*. Thus, the denominator is less than zero. The first term of the numerator is negative and the second term tends to be positive (for normal and superior goods). Purchase of Q tend to rise with income $(Q_Y>0)$ and the effect of income on the effect of price on demand tend to diminish (becomes less negative, e.g. shifts that Q_P upwards, so $Q_{PY}>0$). Thus, the overall sign of the numerator is negative, which given the negative sign of the denominator implies that the overall effect of an increase in income is positive—assuming that income does not affect the marginal cost of searching.

If income were included in the opportunity cost function—which is plausible in many cases, the marginal cost of searching would tend to increase with income. In that case, the term $-C_{NY}$ would be in the numerator, and the overall sign of equation 9.4 would be undetermined without further assumptions. It would depend on whether the purchase effect, the effect characterized in the model as developed, dominates or is dominated by the cost effect.

It is probably the case that millionaires search less regarding the prices of ordinary goods and services than middle class persons do, because for them the opportunity cost effect dominates the purchase effect. However, it is likely that searching by middle-class persons varies quite a bit because of differences in tastes (e. g. differences in tendencies toward prudence or frugality). Thus, the propensity to search for low prices varies among individuals with the same income as well as among individuals with different incomes.

The insight provided by Stigler's model is that the rational choice model can easily be extended to take account of opportunities to learn new things that in at least some cases induce better (less error-prone and more fruitful) decisions to be made and actions undertaken. Moreover, it can explain the distribution of prices observed in real—as opposed to idealized—markets. Knowledge does not have to be perfect, nor given, for markets to operate. Rather, imperfect information can be incorporated into models to provide new insights.

Indeed, assuming that information is "given" is a mistake when analyzing choice settings and/or markets where information or informational differences clearly exist and are likely to be important.

B. Relatively Informed Demand for Information: The Value of Larger Samples for Firms

The simple logic of price searching also applies to firms with respect to their inputs and production processes. Firms have an interest in purchasing their inputs at the lowest price possible—and so they will engage in a search process similar to that described above for inputs— where their demand for those inputs is based on the marginal-revenue product of those factors of production. Similarly, sellers may search for the highest price that a good or service can be sold for. Individuals often sell "used" goods in local markets, where resale values is not always obvious. And, there are cases in which expert opinions are sought for resale value, as with prices of paintings by a highly regarded painter or a house with unusual characteristics.

Most firms sell lots of similar things or services and when their products are a bit unique, the price at which their output can be sold is determined by their demand functions, as developed in chapters 3 and 4. As a consequence, firms are more interested in the nature of their demand functions than in the highest price at which they can sell one unit of their product. Nonetheless, some insight about a small firm's own demand functions can be obtained by simply looking at prevailing prices of other similar goods and services.

However, for larger firms producing products with significant differences, estimates of their demand function help to determine both the optimal scale of their production facilities, the price at which their product or service is likely to be sold (on average) and also the extent to which the product is likely to be profitable. As developed in Chapters 3 and 4, the output and pricing decisions of firms are based partly on their unique demand functions. In Marshallian markets where demand curves are "flat," it is the overall world supply and demand that determine prices—in which case to predict prices firms would have to estimate global supply and demand conditions, which is an even more difficult problem to address. In Ricardian markets with variated products and production methods, reasonably good estimates of demand functions are very useful.

This section focuses on a single firm's effort to estimate its down-ward sloping demand function when more accurate estimates are more costly than less accurate ones—partly because more data are often required for more accurate estimates.

The Usefulness of More Precisely Estimated Demand Functions

The profitability of a firm is affected by the variance of sales, because production normally takes time, and inputs for production are often purchased or contracted for long before a product is brought to market. The higher the variance in estimated demand is, the more storage will be required during periods of lower-than-average demand when prices fall temporarily below production costs. Thus, the higher the variance in sales, the higher are production costs and the more likely a firm experiences excessive inventories of both outputs and inputs. Estimation errors are costly. Accurate demand estimates reduce costs and allow better estimates of profits to be undertaken.

The variance of profit flows—as well as their average—are of interest to firm owners. The usual way that economic textbooks model firms implies that firm owners are all risk neutral, which is to say that they maximize expected (average) profits without worrying about the variance of those profits. However, that is clearly less realistic than it could be. Entrepreneurs and shareholders may be less risk averse than the average person, but surely very few of them are indifferent to risk—because their own wealth and lines of credit are finite. Risk aversion implies that the variability in profits will matter to most owners, financiers, and shareholders.

The variance in profits is partly determined by the accuracy of its estimated demand and cost functions, and partly by irreducible random fluctuations in those functions. Better estimates can not

reduce such fluctuations to zero, but they can reduce profit variability by reducing estimation errors, which in the limit reduces profit variability to that which cannot be avoided. Accurate long run forecasts enable the scale of one's production processes to be properly chosen, which affects both future revenues and costs. Thus, entrepreneurs and firm owners have an interest in accurate estimates of their future demand and cost functions for their planning period. A more accurate estimate (smaller variance) increases risk-adjusted profits by reducing planning errors, storage, and re-contracting costs.

The deterministic models of firm decisionmaking developed in chapter 3 ignored randomness and uncertainty. The extended models developed in chapter 7 illustrated how varability in demand and costs can be taken into account. The same models can be extended to take account of estimation errors in demand by assuming that firm managers and entrepreneurs are uncertain about the values of some of the parameters and variables that determine the precise values of their demand functions. Demand, in principle, may be mechanically causal (as in the models in chapter 2) or stochastically causal (as in the models of chapter 7).⁷

In either case, firms with downward sloping demand functions have to estimate future demand to determine their optimal output and pricing strategies. We have shown how decisions to maximize expected profits can be modeled in chapters 7 and 8. In this section, we'll explore the case in which the variance of profitability can be reduced by collecting data that allows a more precise estimate of a firm's demand and/or cost functions.

The main issue such cases is how much data a profit-maximizing firm should collect and analyze. Conceptually, firm owners maximize utility—just like everyone else—but in a choice setting where their income is determined by their own investment decisions, rather than determined entirely by market forces (supply and demand). Among their investments are investments in data that can improve estimates of future profits. If their utility functions are strictly concave (as assumed in previous chapters), firm owners and entrepreneurs will be risk averse, and invest some of their

⁷ Most of the statistical models that ground econometric studies of demand assume that some degree of randomness exists beyond the core deterministic aspects of demand. That is to say even if one knew exactly what consumer income was and the prices of other goods, the demand function would not be completely identified, because of random "shocks" of one kind or another beyond the variables of the demand function, as with $Q^{D} = a + bY - cP + e$, where a, b, and c are constants (parameters of the demand function), Y is consumer income, P is the selling price of the good, and e is a purely random factor with mean zero and variance σ^{2} .

wealth in data collection to reduce their risks from fluctuations in their income associated with profit variability. Unfortunately, writing this optimization problem down in general terms does not shed much light on such choices, although it can characterize such choices does shed some light on the factors that determine such choices, as shown below in a footnote.⁸

The Optimal Precision of a Firm's Estimated Demand Functions

An alternative way to get at the effect that more data and analysis has on risk is to use another idea from Chapter 7, namely the idea of a risk premium. To undertake a risky investment a risk-averse investor requires a risk premium—a profit above that associated with the risk-free (low risk) options that he or she is aware of. This idea can be used to characterize risk adjusted expected profits by subtracting an entrepreneur/owner's risk premium from expected profits. The risk premium demanded by a risk-averse investor rises with the variance of the expected profits. That variance can be reduced by collecting and analyzing more data, although usually not to zero.

⁸ This footnote illustrates how that can be done. Suppose that $f(\Pi)$ is the frequency distribution of profits (Π) and $f(\Pi | D)$ is the conditional distribution of profits for given amounts of data collection and analysis (D). Combining ideas from chapter 2 and chapter 7 allows the expected utility maximizing choice for a two good utility function with uncertain income (here Π) to be written as $U^e = \int u(X_1, (\Pi - P_1X_1 - P_DD)/P_2))f(\Pi | D)$. This formulation uses the expected formula for a variable with a continuous frequency distribution. Π is the individual owner's or entreprenur's share of the firm's profits. X₁ is the quantity of good 1. P₁ is the price of good 1, and P₂ is the price of good 2. As, written, this firm owner/entrepreneur has two control variables, X₁ and D, assuming that the firm of interest produces a third good or service not purchased by the firm owner/entrepreneur.

Differentiating with respect to X_1 and D yields two first-order conditions that characterize the entrepreneur's consumption of goods 1 and 2, and his or her investment in data purchases and analysis.

$$U^{e}_{X_{1}} = \int [u_{X_{1}} - u_{X_{2}}(P_{1}/P_{2})]f(\Pi|D) = 0 \quad \text{at} \quad X_{1}^{*}$$
$$U^{e}_{D} = \int u_{X_{2}}\left(\frac{P_{D}}{P_{2}}\right) + u(X_{1},(\Pi - P_{1}X_{1})/P_{2}))f_{D} = 0 \quad \text{at} \quad D^{*}$$

Notice that the first first-order condition is satisfied where $u_{X_1} - u_{X_2}(P_1/P_2) = 0$, as in the nonstochastic version of this choice problem—although in this case that choice is affected by variation in the individual's profits from producing another good or service. The second first-order condition demonstrates that data and analysis are selected to maximize expected utility, in this case by setting the marginal opportunity cost of data collection and analysis (the first term) equal to the increase in expected utility generated by additional data and analysis. That increase in expected utility is generated by reducing the variance of the estimated distribution of profits. Both first-order conditions are simultaneously satisfied at the expected utility maximizing choice. Suppose the firm knows that its inverse demand function is P = p(Q, Y) where Y is the average income of its consumers and Q is its output, as in Chapter 3. Suppose that Y is randomly distributed with mean Y^e and variance σ^2 , where both Y^e and variance σ^2 have to be estimated by the firm. The variance of both estimates is assumed to fall as more data are gathered and analyzed, as implied by standard statistical analysis. The real underlying variance of consumer income, of course, is not affected by data collection analysis, but the variance of estimated consumer income is. Consumer income may still vary—as for, example, because of business cycles. The estimation error is normally reduced through additional data and analysis, D, (albeit with diminishing marginal returns).

The risk premium of the investment of interest (here the production of goods for sales in the future) rises as the overall variance (estimation errors plus underlying randomness) increase. Thus, the risk premium (here a decrement) for choosing output Q can be written as $R = r(D, \sigma^2)Q$. That premium (decrement) tends to fall as more analysis, D, is undertaken—although not to zero because the true underlying risk, σ^2 , is not affected by the research and estimation process.

Given this characterization of the choice setting, the expected risk adjusted profit can be written as the expected profit minus the risk premium:

$$\Pi^{e} = p(Q, Y^{e})Q - c(Q, D) - r(D, \sigma^{2})Q$$
(9.5)

Differentiating with respect to Q and D yields two first order conditions.

$$\Pi^{e}{}_{Q} = p_{Q}Q + p(Q, Y^{e}) - c_{Q} - r(D, \sigma^{2}) = 0$$
(9.6)

$$\Pi^{e}{}_{D} = -c_{D} - r_{D}Q = 0 \tag{9.7}$$

Both first-order conditions have to be satisfied simultaneously, but some insights can be obtained by looking at them separately.

First, equation 9.6 implies that the greater that risk (e.g. risk premium) the lower output tends to be. A risk-averse entrepreneur will produce somewhat less than the amount that maximizes expected profits. How much less will vary with his or her degree of risk aversion and the underlying variance of the phenomena (here average consumer income, Y^e) of interest. Second, equation 9.7 implies that the extent of the research to estimate the probability distribution of the risky variable (average consumer income, Y^e) tends to fall as marginal cost increases, c_D , and to rise the more rapidly research reduces estimation errors (which reduces the marginal risk premium). There are benefits associated with greater data collection and analysis, although they are limited by the extent to which estimation errors are reduced and by the marginal cost of data collection and analysis, c_D .

Note that the term, $r_D Q$, is the source of a large firm's demand for economists and/or for its use of external consulting firms that employ economists. It is, thus, the effect on risk adjusted profits that accounts for most of the non-academic private-sector job market for economists.

III. Ignorance

The sampling approach to imperfect information is a plausible characterization of a variety of informational problems—namely, those in which one has an understanding of the phenomena of interest (as with prices or demand functions) but is less certain about the actual properties of the distribution of possibilities than possible with more information and analysis. This approach assumes that much is already known to the relevant decision maker—the domain of possibilities and often the type of probability distribution associated with the stochastic phenomenon of interest. More information allows one to better understand the phenomenon of interest (lowest price, demand function, etc.) but does not uncover entirely new phenomena or possibilities previously unknown to the decision maker of interest (consumer, firm, scientist).

Ignorance, in contrast, occurs when relevant details about a phenomenon are entirely unknown. Indeed, the entire phenomenon may be unknown. For example, the possibility of telegraphs and telephones for long distance communication prior to 1800, a low cost method of refining aluminum before the Hall process was worked out, or such innovations as electric lightbulbs, washing machines, automobiles, heavier than air passenger flights, radio, television, solid state electronics, personal computers, cell phones, and satellite internet—to name just a few of the surprising innovations that created unexpected new markets for new products and services. Instances of contemporary ignorance include our limited knowledge the foreign languages (many of which we do not know exist), the manner in which computers are constructed and operate, and the distribution of talents and knowledge among persons that we have never met.

In some cases, ignorance can be said to be chosen—e.g. rational ignorance—because most of us know that unknown things exist (e.g. unknown languages or fields of science) yet take no steps to reduce our ignorance (by consulting lists of languages and learning a few words in each one), or by our failure to spend even a bit of time to fill gaps in our knowledge of science or engineering (by reading books and journals or by taking courses on them). In some cases, ignorance is natural, a human phenomenon. There are relationships and things that will be known in the future are not yet known. In others, it is a personal matter—some people have mastered languages and fields of science and engineering that we are unaware of, but not everyone else is unaware of.

Both categories of ignorance are relevant for economics. The natural ignorance of humanity creates opportunities for significant innovations to occur. The natural ignorance of individuals creates demands for expert advice and other specialized services. Both would be impossible if we all knew everything that is possible—and books such as this one, as well as universities, would be unnecessary—indeed, a complete waste of time.

A. Ignorance. Latent Demand, and Innovation

Ignorance is relevant for the theory of demand for several reasons. One is that ignorance creates the possibility of latent demands—demands that individuals have but are unaware of, and which become actualized when a new product is brought to market and seen for the first time or learned about through the mass media or one's social networks. Several of the products that can be regarded as Schumpeterian innovations (e.g. consequential innovations) appear to be demands that were latent. A few entrepreneurs believed that particular latent demands existed, and attempted to profitably serve them. The automobile is one such product. The smart (cell) phone is another—but there are many dozens of such products brought to market every year—not all of which are successful, of course, which demonstrates that latent demands do not always exist for every new product imagined by an entrepreneur.

Ignorance makes mistakes possible—indeed likely. Nonetheless, changes in the product mix brought to market through time clearly demonstrate that many instances of latent demands have existed in the past.⁹

⁹ In my own case, I remember learning about a number of new products that I did not know existed when I attended college as a young man. For example, my hometown did not have any highend stereo shops, and I learned of the possibility of high-quality reproduction of sound both from previously unknown fellow students who became friends and who had such systems and also from the local stereo shop. The latter also taught me about higher quality sound sources than I had known about. (Similarly, it was a surprise to me that not all albums were equally well recorded in those days.) Higher quality reproduction systems involved many higher quality products that jointly more faithfully reproduced the sounds from vinyl discs and CDs(phonographs, cartridges, CD-players, preamps, power amps, and speakers). Who knew? Obviously many others did or there would not be

Modeling Latent Demand

Such demands can be easily modelled, if we drop the assumption that individuals fully understand or know their own tastes and utility functions.

The effect is easiest to demonstrate with concrete three-good utility functions. We'll assume that Al's true utility function is $U=A^{a}B^{b}C^{c}$ and that Al has W dollars to spend. The prices of the three goods are P_{a} , P_{b} , and P_{c} . However, initially Al does not know that good C exists and so acts as if her Utility function is $U=A^{a}B^{b}$ —which is what it would have been if c=0. (Recall from high school algebra that $C^{0} = 1$.) Forming a Lagrange function for the pre-knowledge of C choice, we have:

$$L = A^a B^b - \lambda (W - P^a A - P^b B)$$
(9.8)

Differentiating with respect to A and B, setting the results equal to zero and a bit of algebra (see that worked out in Chapter 2) generates the following two demand functions for Al.

$$A^* = aW/P^a(a+b) \tag{9.9}$$

$$B^* = bW/P^b(a+b) \tag{9.10}$$

If for example, a=b, Al would spend half of his/her money on each good, with the actual quantities purchased varying with the price of the good or service.

Now consider Al's optimization problem after good C has been discovered and exponent c has been found to be different from zero. The Lagrange function becomes:

$$L = A^{a}B^{b}C^{c} - \lambda(W - P^{a}A - P^{b}B - P^{c}C)$$
(9.11)

Differentiating with respect to A, B, and C, setting the results equal to zero and a bit of algebra now produces three demand functions—each with a similar functional form.

$$A^* = aW/P^a(a+b+c) \tag{9.12}$$

$$B^* = bW/P^b(a+b+c)$$
 (9.13)

$$C^* = cW/P^c(a+b+c)$$
 (9.14)

such shops, but not me. My latent demand was activated, and I began aspiring to purchase, and eventually purchasing, relatively high-end (accurate) stereo equipment. Like many college students, many of my long-standing hobbies were developed while a student. Other later latent demands revealed by innovation, for me at least, include lap tops, jet travel, the internet, and cell phones.

Notice that the denominator is larger than it was in the previous case, so the amounts spent on goods A and B falls, and that decrement is now spent on good C. In this manner, the effect of activating a previously latent demand is to shift expenditures around in one's personal budget—reducing expenditures on many previously known goods and services so that some of one's wealth or income can be spent on the "new" good—at least goods that are new to Al.¹⁰

In extreme cases---although not ones associated with this concrete form of a utility function—expenditures on some of the preexisting goods may fall to zero.

Note that in either case, this model illustrates one way that a commercial society's impulse for creative destruction can be modeled. The successful invention and marketing of a new product upsets the pre-existing equilibrium—here through effects on demand for goods and services generally reducing demands for pre-existing goods while increasing it for goods that previously had only latent demands..

B. Rational Ignorance and Specialization

Adam Smith (xxxx), George Stigler (xxxx), and James Buchanan (xxxx) all have argued that that much of the efficiency of production in markets is generated by specialization—by which they mean production by organizations and individuals that produce a narrow range of goods and services for sale in markets. That narrowness in some cases reduces the satisfaction associated with work itself, but can substantially increase the outputs of firms and individuals relative to that which could be produced by "well rounded" organizations and individual "jacks of all trades." Hayek (xxxx), in turn, suggests that specialization itself—although motivated by market incentives—is largely informational in nature. It arises because of differences in the productive knowledge acquired and internalized by individuals over the course of their lifetimes (e.g. their "skills").

Individuals know quite different things—partly because of differences in talents, but mainly because of different choices made about which kinds of information to acquire through experience, reading, and education. It these differences that create comparative and absolute advantages in production, and it is the demand for the things and services that can be produced that account for

¹⁰ Here is should be noted that reductions in spending on all other goods does not necessarily occur with different utility function than the multiplicative exponential function used here—but the overall expenditure on previously known goods has to fall if new expenditures on the good learned about are to be undertaken. Expenditures on some previously purchased goods may actually increase (e.g. complements to the new good) while most others fall.

the pattern of market rewards for different kinds of market-relevant productive knowledge. Market prices, in turn, are affected through specialization's effects on expected future income, which encourage individuals to specialize.

That rates of return among skill sets differ (e.g. are associated with different wages and/or job satisfaction) generate demands for the information sets associated with those skills—e.g. demands for training, experience, and education. Those rewards often induce "corner solutions" in individual demands for information—which is to say, they induce many individuals to devote essentially all of their time and attention on the subset of information relevant for their chosen speciality, rather than encouraging one to divide one's time and attention among a broad variety of types of information.

That is to say, decisions to pursuit a career often induce individuals (and organizations) to remain rationally ignorant about most types of information that are irrelevant for their speciality. For example, if one pursues a degree in economics, one usually will forgo learning about accounting, physics, electrical engineering, and so forth because the time and attention required to obtain the training necessary to qualify as an economist precludes such investments.

In contemporary markets some forms of "narrowness" yield much higher salaries than a broader more general education normally does.

A Model of Career Choice

The conscious demand for career-specific information tends to occur after individuals have acquired a good deal of information over many years. For example, most students begin to think about careers in high school or college—or later. As young adults, they are no longer naturally ignorant about all topics. Most have some information about career opportunities—although their knowledge of possibilities is often quite incomplete.¹¹ Obviously, one cannot choose a career that one is unaware of. Thus, our first model of career choices (e.g. specialization) involves a decision between two careers about which at least a bit is known.

¹¹ In my own case, after a middle-class background, I had no idea what economics as a field was when I went off to college, and I took a college course in economics only after the strong recommendation of a close friend although I still had little idea about what would be learned. I had remained naturally ignorant about economics as a field.

The decision of interest is time allocation choice—where the time is devoted to acquiring skills (knowledge) of one kind or another. In the case of college education and other forms of technical education, personal expenditures may also be involved, but the time allocation problem is sufficient to illustrate some of the factors that influence decisions to specialize.

Suppose that Al is choosing which of two careers to focus on, and leisure. The careers in turn are expected to produce various combinations of income and job satisfaction. The exact relationship between investing in a career (studying particular subjects) and income and satisfaction are not known, but enough information is possessed that statistical model of possible outcomes are used for Al's choice of a speciality. We'll first model choices when specializing is the only possibility, and then consider a choice where one can choose between complete specialization and various combinations of those specialities (as with double majors) are possible. Musicians, for example, often pursue two careers—a "day job" and a "night job" (the latter being their more enjoyable music performances).

In both cases, a forward looking individual will imagine the expected utility associated with each of the possible career paths, which is partly a matter of expected income, partly of job satisfaction, and partly of the leisure sacrificed to qualify for positions in the speciality of interest. In many cases, decisions to specialize are once in a lifetime choices, but they may also be regarded as choices made with a shorter finite time horizon—a plan for the next several years. Many complexities can be taken into account, but as with most models, the aim here is to isolate some essential features of career choice—e.g. decisions to specialize—that are likely to be factors in most such decisions, rather than all of them.

Choosing Among Specialities

To understand some of the essential aspects of decision's to specialize, we'll consider a choice setting in which there are just two specialities, A and B, that are considered by AL and that Al's planning horizon consists of two periods—a training period and a finite income-earning period of employment in his or her speciality. We'll also assume that there is no prospect of failure during the training period as long as one invests the required time in the training and that the distribution of income in each of the fields is bimodal, with the probability of high-income being P^A for career A and P^B for career B. The training periods differ, with T^A and T^B being the periods of study, and S_A and S_B , being the time requirement for study during the period in which information is acquired and internalized (becomes part of one's personal knowledge). After certification, the high-income

associated with speciality A earns A^H and the lower income A specialists earns A^L. The probability of being in the high-income group of occupation A is P^{AH}. Similarly, the high-income group of occupation B earn B^H and the low-income group, B^L, with the probability of being in the high-income group being P^{BH}. Each person allocates D hours each day between leisure and study or between leisure and work. In the work period, individuals are assumed to be free to pick the length of their work years (although that is not, of course, always true), although not in the training period. Utility is assumed be generated by Leisure (L), Job Satisfaction (J), and material comforts associated with income (Y).

Expected utility realized during the planning period, F, can be written as:

$$U^{A} = \int_{0}^{T^{A}} u(D - S^{A}, 0, 0, t) dt + P^{AH} \int_{T^{A}}^{F} u(D - W^{A}, W^{A}, A^{H}W^{A}, t) dt + (1 - P^{AH}) \int_{T^{A}}^{F} u(L - W^{A}, W^{A}, A^{L}W^{A}, t) dt$$
(9.15)

for career A, and as:

$$U^{B} = \int_{0}^{T^{B}} u(D - S^{B}, 0, 0, t) dt + P^{BH} \int_{T^{B}}^{F} u(D - W^{B}, W^{B}, B^{H}W^{B}, t) dt + (1 - P^{BH}) \int_{T^{B}}^{F} u(L - W^{B}, W^{B}, B^{L}W^{B}, t) dt$$
(9.16)

for career B.

Notice that although many aspects of career decisions are taken into account in this model, there are just two control variables for Al. He or she can choose a speciality and decide how hard to work in the post training period. The career choice in this case is entirely determined by the higher of the two expected utility functions—which is partly a consequence of calculus (determining the ideal work rate in each career), although specialization is not a choice that calculus sheds direct light on. Al simply chooses the occupation (speciality) that generates the highest expected utility, given how hard he or she expects to work in the training and post-training periods and the income and job satisfaction anticipated for the two careers under consideration.

The first-order condition that characterizes Al's post training work effort are notationally similar but differ because of the parameters of the post-training work periods, including differences in the wage rates. Al's work effort in the post-training period if career A is characterized by:

$$P^{AH} \int_{T^A}^F -u_L + u_J + u_Y(A^H) dt + (1 - P^{AH}) \int_{T^A}^F -u_L + u_J + u_Y(A^L) dt = 0$$
(9.17)

And that associated with career B is:

$$P^{BH} \int_{T^B}^{F} -u_L + u_J + u_Y(B^H) dt + (1 - P^{BH}) \int_{T^B}^{F} -u_L + u_J + u_Y(B^L) dt = 0$$
(9.18)

In both careers, the anticipated work effort varies with the expected present value of the marginal opportunity cost of working more (reductions in utility from diminished leisure) and the marginal returns from working (including both additional job satisfaction and material comfort).

The effects of the differences among careers (training period, required study time, odds of being among the high versus low-income graduates, job satisfaction, and the high and low wage rates) can all be characterized using the implicit function differentiation rule. And through their effects on expected total utility in the planning period (using the envelop theorem), their effects on career choices can be analyzed and discussed—although only in cases in which Al is initially indifferent between the two career paths would small changes in any of the parameters of the choice setting affect his or her choice.

Even without the implicit function derivatives, the utility functions and first-order conditions provide considerable insight into how the various factors affect an individual's choice of speciality.

When the wage distributions are not very different—differences in job satisfaction are likely to be decisive. When job satisfaction and training costs are approximately the same in the two specialities, then expected wages (average wages) and the range of wages (difference between the high and low wage subgroups of the specialities) will determine choices. If average wages and job satisfaction are the same, then the variance within a speciality will determine the choice if Al is risk averse. In that case, the speciality with the least variation in wage rates will be preferred to ones with more variation. To accept the risk associated with a greater variation in wages, individuals would require a risk premium—higher average wage—to compensate for the riskiness of the speciality.

If job satisfaction and the wage distribution are similar, then the speciality with the lowest training cost will be selected. Similarly, the length of time that must be devoted to training before one can begin earning a salary will be important for persons with positive rates of time discount (which is represented as the effect of t on utility in Al's utility function). For a person with a high discount rate, the shorter the training period the better.

Choosing to Specialize or Not

Modest extensions of the model can also be used to think about why individuals choose to specialize rather than to divide their time between the two fields under consideration. The joint choice can be modelled simply by changing the time constraint so that time is allocated between two types of work and combining the two expected utility functions after deciding which training period (major or advanced degree) is completed first. We'll assume that what might be called certification matters, and that to qualify for either specialty requires spending a given time in training—as in the first model. We'll also assume that training for speciality A takes place first, followed immediately by training for speciality B. Given the probabilistic and temporal aspects of the model, the new optimization problem looks quite complicated. Maximize:

$$U = \int_{0}^{T^{A}} u(D - S^{A}, 0, 0, t) dt + P^{AH} + \int_{T^{A}}^{T^{B} + T^{A}} u(D - S^{B}, 0, 0, t) dt + P^{AH} \int_{T^{B} + T^{A}}^{F} u(D - W^{A} - W^{B}, W^{A}, A^{H}W^{A}, t) dt + (1 - P^{AH}) \int_{T^{B} + T^{A}}^{F} u(L - W^{A} - W^{B}, W^{A}, A^{L}W^{A}, t) dt + P^{BH} \int_{T^{B} + T^{A}}^{F} u(D - W^{A} - W^{B}, W^{B}, B^{H}W^{B}, t) dt + (1 - P^{BH}) \int_{T^{B} + T^{A}}^{F} u(L - W^{A} - W^{B}, W^{B}, B^{H}W^{B}, t) dt + (1 - P^{BH}) \int_{T^{B} + T^{A}}^{F} u(L - W^{A} - W^{B}, W^{B}, B^{H}W^{B}, t) dt$$

In this case there are two first order conditions—one for W^A and one for W^B , rather than one since a dual career is contemplated. Mathematically, the first order conditions closely resemble those computed in the previous section except that the starting time of the employment period which is now $T^B + T^A$, rather than T^B or T^A because following both careers requires that Al meet the qualifications for each. Substituting the optimal values for W^A and W^B into equation 9.19 allows the expected utility of the two-career future to be characterized.

The choice about specializing requires comparing the expected utility associated with the idealized two-career future to two specialized career paths characterized in the previous section. Specialization is the rational choice whenever one of the single-career futures is expected to generate more satisfaction (utility) than the other two possibilities.

Normally, in neoclassical economic models, "corner solution" are neglected and "interior solutions" are the main focus, as with the focus of consumer choice models on the middle points of their budget constraints rather than their end points. That is not always true of consumer or production models for reasons that we have noted in passing in earlier chapters (the most common quantity purchased by a consumer in a grocery store is zero), but most people spend their incomes on many different goods and so interior solutions are the ones that characterize the most relevant aspects of the spending decision—e.g. the variety of goods and services actually purchased.

However, in the case of specialization, $W^i=0$ is very commonplace for all but a single type of speciality. Most people work at one job within one specialty. There are a huge number of occupations that persons could have devoted a little time to, than they actually do. Most college students chose only a single speciality during their time in undergraduate and graduate educational programs while at a university and also during time in a trade schools. So, with respect to specialization, corner solutions are important (e.g. solution in which either W^A or $W^B = 0$.

This implies that the combination of job satisfaction and wages associated with mastering a speciality are generally anticipated to be well above those associated with double or triple majors for most students—and also with respect to unmodelled generalist degrees.

If the expected utility functions after all the substitutions are strictly concave and there are no certification requirements for employment, then the usual geometry of optimization would imply that being a jack of all trades would be the best strategy for most individuals, including our representative individual, Al. To induce specialization, there must be some non-concave aspects to the decision problem. In this case, evidently the wage rates and/or job satisfaction are much higher for specialists than non-specialists—other things being equal—or the training cost for more than one speciality must be very time consuming and challenging for most individuals. It may be easier to master a narrow set of complementary information than a broader range of information that lacks complementarities.

In fields where training is for short periods and does not materially increase salaries or job satisfaction—less specialization takes place and persons can move among unspecialized jobs (jobs that can be easily and quickly mastered) but not from such fields to the more specialized and higher paying fields. Among economists, such unspecialized careers are often referred to as "unskilled" position, even some training nearly always necessary and a person without that training would be

unable to perform the duties of such occupations. Even mopping floors and washing windows requires some training to do them well.

Entry Barriers and the Unequal Returns from Different Specializations

That wage rates among specialities differ for long periods of time is also of interest. As people choose among specialities based on wage rates and job satisfaction, rates of returns to particular specialities should equalize at the margin—in which case specialization may still pay, but what one specializes in should not matter very much. However, in many cases, such equalization of the returns from specialized education does not occur. Doctors and lawyers, for example, have for nearly a century realized average salaries that are much higher than other specialities involve rough equal periods of training—such as college professors—and wage rates have varied significantly among the specialities pursued by doctors and college professors—rather than equalizing as might have been expected.

For such unequal returns equilibria to occur, there must be barriers to entry of some kind. One such barrier would be the scarcity of particular talents necessary to obtain the speciality of interest. For example, short persons almost never are hired as professional basketball players because height by itself tends to make basketball players more productive as players, even if they have spent equal hours training as athletes in this speciality. Especially tall and athletic persons are only quite rare, and there for markets across sports specialities do not fully equalize. Similarly, it is possible that particular types of minds are required for particular professions—for example memory may be more important for some fields than others. The scarcity of particular types of talents is one such barrier to entry in specialities, a natural one. Variation in family-based training is another—as child prodigies at piano playing often grow up in musical families—e.g. families of musicians where training is freely available, and instruments are routinely provided to family members at a young age. Another can be various cartel-like strategies that limit opportunities for people that are sufficient talented to specialize in a particular field but are prevented for one reason or another from obtaining the training necessary for a particular speciality.

What is most important for this section of Chapter 9, however, are not the variety of entry barriers that cause rates of return in human capital to vary widely among occupations, but that specialization often makes sense from the perspective of individuals and organizations.

For this to be so, corner solutions in many information markets (e.g. human capital markets) must be commonplace. This implies that there are increasing returns (increasing marginal products) associated with some forms of information collection and mastery, and that therefore increasingly high marginal revenue products are associated with an important subset of narrow and intensive training programs (as with college majors) than with broader more generalized training programs. Sufficiently so, that the risk-reducing advantages of being a generalist—being a renaissance man or woman—are dominated for a large fraction of a contemporary economy's workers by the extraordinary returns achieved bh focusing one's attention on a narrow domain of productivity enhancing information—while remaining rationally ignorant of a wide variety of useful information in most other fields of knowledge. Expertise often pays.

IV. The Importance of Prices in a Society of Weakly Informed Specialists